

Communism is not yet dead

Reports that Marxist-Leninism is dead are exaggerated, even dangerous. Mr Mikhail Gorbachev may be on the way to shedding some of the party's power, but he still needs its help.

MR Mikhail Gorbachev's triumph in the Central Committee of the Soviet Communist Party last week has been widely taken in the West to mark the end of communism in the old model of Marx and Lenin, but that is a mistake, even a serious miscalculation. What Gorbachev has done is to win grudging approval for the abolition of Article VI of the Soviet Constitution, which guarantees the Soviet Communist Party a "leading role" in the conduct of national affairs. So long as events in the Soviet Union do not move as quickly as they have done in Eastern Europe during the past few months, the Supreme Soviet called for June will no doubt then approve a new constitution. Meanwhile, the Russian Republic will be holding elections early next month at which party candidates at all levels of the administration are likely to take a further drubbing. Yet there are two respects in which none of this implies that communism is dead.

First, the party apparatus remains in being, with its complicated hierarchy of salaried secretaries and officials at whose pinnacle is General Secretary Gorbachev himself. It is natural enough that many of those concerned fear (and thus resist) the ending of their incestuous power, which has enabled them since Lenin's death in 1923 to influence the appointment of public officials of all kinds — and then to determine their behaviour. (The reforms of 1987 requiring laboratory directors and some other officials to be elected by their staffs, of dubious merit and at best cumbersome, were at least a way of taking important positions out of party hands.) The snag now, for reformers like Gorbachev, is that there is no other way by which essential appointments can be made. While the government pretends to run everything, or nearly everything, it needs the party apparatus as a kind of civil service, with all the nepotism and influence-peddling that has become traditional.

Hope

The best hope, for the months ahead, is that the reform-minded members of the party will be able to restrain the worst excesses of the local and regional party chiefs. The greatest danger is that the party, sensing the permanent decline of its power, may seek to exploit the turmoil in the Soviet Union now. Gorbachev has been tormented in the past few months by outbreaks of nationalism in the peripheral republics, but an appeal to Russian nationalism by an influential section of the party would be dangerous

indeed. That is yet another reason for hoping that it will hang together through the difficult transition that lies ahead.

It would also be a grave mistake for outsiders to overlook the temper of Soviet society, at least in its heartland of the Russian Republic. For demographic reasons, there are few among the adult population of the Soviet Union who can recall the years before the revolution in 1917. Stalin's tyranny notwithstanding, patriotism and past illusions of progress have engendered a powerful sense within the Soviet Union that social justice is inseparable from the enlightened practice of Marxist principles. These, from whose ranks the party reformers must come, will be as important to Gorbachev in the next few years (especially months) as the more radical reformers. They will not disappear altogether.

Meanwhile, Gorbachev, or someone of his stature not now in sight, has to make the Soviet Union work as if it were a modern state. The disappearance of Article VI in June will count for little if there is not by then a plan for turning a large chunk of industry and the whole of the farm economy into autonomous entities, free from the shackles of central planning. That does not mean selling off industry or farms to those with the most rubles in their hands, or even forsaking communism for capitalism, but merely that there should be objective mechanisms that decide which enterprises survive and which go out of business. It will be time enough to worry about the long-term future if the next winter can be made less chaotic than that now ending. □

Embryos win rights

The British Embryo Bill has survived the first hurdle, but its unintended consequences remain to be discovered.

PARADOXICALLY, some British researchers are relieved that the prospect of being thrown into jail on account of the research they do has become significantly more certain. The House of Lords last week approved in principle the British government's Embryo Bill, whose most controversial provision (so far) is that research with human embryos should be allowed up to 14 days after fertilization, but only with the approval of a statutory authority yet to be created. Unlicensed research, or any



research after the 14-day limit, will be a criminal offence, punishable with a jail sentence of up to two years. So why the relief? Because there had seemed a danger that the House of Lords would choose the other option offered by the government, and opt for a ban on all research with embryos.

The Embryo Bill gives effect to the recommendations in 1984 of the Warnock Committee on embryo research, the practice of *in vitro* fertilization (IVF) and related matters. The bill, of course, is shockingly late. When it is generally acknowledged that public anxiety will be quietened only if people know what researchers are doing and understand why, regulation has fallen to a committee organized by the Medical Research Council and the Royal College of Obstetricians and Gynaecologists whose intervention rests on the voluntary compliance of researchers and practitioners. The committee has done an excellent job, but will also be relieved that the end of its tenure of a non-statutory office is in sight. But even now, the issue is not settled. The House of Lords still has to pick over the details. Only then, a couple of months from now, will the House of Commons have its say.

Delight at the House of Lords decision notwithstanding, it is relevant that the bill as it stands is deficient in many ways, not least because of the 14-day limit on research with human embryos. Not that there is at present any great push for laboratory work that would breach the limit. Most proposals to the voluntary committee are intended to improve the effectiveness of IVF for the treatment of infertility. Otherwise, for the general understanding of early embryonic development, people seem content to work with mice or other species. But the time will come when Warnock's arbitrary choice of 14 days, an estimate of the appearance in the developing embryo of the primitive streak from which nerve tissue emerges, will impede research of potential value. It would have been better to have provided the licensing authority with more explicit statutory guidelines and then to have done without a limit.

There is a wider reason for regretting this feature of the bill. The 14-day limit will be generally understood to offer absolute protection to embryos older than 14 days. But if embryos are protected after 14 days, why should it be permissible to abort fetuses after 28 weeks of gestation (the present time-limit, likely to be reduced)?

Dr John Habgood, the Anglican Archbishop of York, did a manful job in last week's debate by arguing that the moral questions that apply to embryo research and abortion are distinct. Nobody can deny that. Research with embryos promises understanding that will be generally beneficial, abortion (under the present British law) protects the health and social well-being of individual women — large numbers of them every year. The obvious difficulty is that this will not be generally appreciated — and that, when it is, the moral case for embryo research will seem the stronger. In short, whatever happens to the Embryo Bill, the 14-day limit is a dangerous hostage given to fortune. □

Animals and friends

The US Congress should avoid passing unworkable animal rights legislation.

THE US Congress seems to be awakening to the need that something should be done about the animal rights movement, but only slowly (see page 580). The diffidence is understandable. Over the past decade, advocates of animal rights have won a dubious but formidable reputation for implacable ferocity. In the circumstances, it is remarkable that three members of the Congress have been brave enough to put their heads above the parapet, introducing bills directed against those who perpetrate crimes against laboratories and against those who work in them. Whether the zealots responsible will be deterred by knowing that breaking into federally supported laboratories may in future be a federal felony, or that they may have the FBI on their tails, is another matter. The stock-in-trade of those that perpetrate violence on behalf of animals is the determination not to compromise with reality.

That is why the Congress would be better advised to concentrate its energies on more manageable objectives. For the past five years, the Animal Welfare Act has been lumbered with a welter of provisions whose full implications are not yet clear — except that they cost money — but which may be in part unworkable. That people embarking on experiments with animals should be required by the act to justify what they intend to an independent panel is not unreasonable. In Britain, researchers have been required for two years to win approval from a government committee (just reconstituted, with Lord Nathan as chairman) and appear not to find the process over-irksome. But in the present climate, the US legislation almost guarantees that the committees become vehicles for rehearsing irreconcilable opinions. The objectives of animal welfare, meanwhile, are over-ambitiously defined. The act needs urgent attention to be made workable.

For the rest, very little can be expected of the Congress. The US administration, the most generous supporter of biomedical research, could be more outspoken in defence of its own spending programme, but there are few votes in that.

Much the most serious burden rests with the scientific community, which should concentrate on one central task: persuading moderate opinion that the use made of animals in research is responsible, humane and directed at goals considered by the generality to be worthwhile. But what purpose will be served by placating moderate opinion, when the men and women with crowbars and spray-cans are incorrigible? Because the wild men, despite the incoherence of their cause, derive their strength from the indifference of moderate opinion to the harm that will be done if biomedical research is made impossible. □

AIDS in children adds to Romania's troubles

- Ceaucescu legacy of public health problems
- Microtransfusions blamed for outbreak

Bucharest

AN alarming outbreak of AIDS among young children in Romania appears to have spread through the practice of giving 'microtransfusions' of whole blood or plasma to malnourished or sick children in order to help build up their strength. The practice is not used in the West.

Five hundred and fifty Romanian children have been identified who are infected with HIV, most of them under three years of age. More than one child in three was found to be infected in tests conducted in hospitals and orphanages in the districts of Bucharest, Giurgiu and the Black Sea port of Constanta. But tests of 280 children in the central and western Romanian cities of Sibiu and Timisoara revealed no infection.

All the children infected before the age of three are likely to die of AIDS-related infections like lymphoma or tuberculosis within two years, according to Ludovic Paun of the Romanian national HIV reference centre. The children are thought to have been infected since mid-1989; Romania had reported to the World Health Organization (WHO) 74 cases of AIDS up to September 1989, including 50 cases in children.

Whether the disease was spread in other ways is still unknown. Young children may also have been infected through the use of unsterilized needles and syringes for vaccinations or vitamin D injections — both are commonly given to children under three years of age in Romania. Sterile equipment was not available in sufficient quantities in Romania and the hard currency was not provided by the government to purchase it.

The WHO sent a team to Romania this week to advise the government on the establishment of a national AIDS programme. David Heymann, an epidemiologist at WHO in Geneva said the new officials in the Ministry of Health were very interested in the programme and would be able to coordinate it capably.

The attitude of the current government towards AIDS stands in stark contrast to that of the Ceaucescu regime. The old regime hindered researchers in their efforts to screen for the HIV virus and prevented researchers from meeting to discuss the problem. Physicians had to use other diagnoses such as "septicaemia" and "unspecified infection" in order to bring the children to a central clinic in Bucharest for treatment.

The widespread use of microtransfusion

was a result of the conditions caused by Ceaucescu's efforts to increase the population of Romania. His regime forbade both birth control and abortion for women who had not yet had five children or reached the age of 45 years. But in many families, there was not enough money to feed or clothe all the children properly and many children were abandoned.

When a child became ill or undernourished, parents often thought that the child would be better cared for in hospital, where, they believed, there would at least there would be warmth, food and medication, said Stiri Brindusa, a medical resident in Bucharest. But often the doctors had few supplies, which led to the use of microtransfusion. The practice has now been stopped.

The Paris-based Medecins du Monde has helped deliver one million disposable syringes, 5,000 sets of reagents for the ELISA test for HIV, and other materials.

Heymann said that screening of all blood and blood donors (of whom 600,000 give blood annually in Romania) would cost millions of dollars. He emphasized that only by pooling the resources of all the aid organizations with government aid from other countries could the epidemic be brought under control in Romania. Several other East European countries have also asked WHO for help in evaluating their national AIDS programmes and will attend a meeting on 27 February to discuss their needs.

Brindusa, who spent much of the past year caring for children with AIDS, said that the turnaround in AIDS policy was for her the "second best effect of the revolution. Now we can say that the disease exists and do something about it." The best effect of the revolution, she said, is "That idiot [Ceaucescu] is dead."

Steven Dickman

Correction: HERA

THE headline of last week's article about the German and US particle accelerators (HERA and SSC) erred in representing the extra cost at Hamburg as \$300 million, which is the estimated extra cost of SSC. The text explains that an extra DM30 million may be spent on improvements to the injector, not the "magnets", at HERA. The article also said that there is "a certain minimum luminosity" below which colliders such as HERA yield no collisions; in reality, the collision rate increases with the intensity of each beam. *Nature* regrets these errors. □

HUMAN GENOME

Keeping them guessing

Paris

THREE hundred scientists from 35 countries met at the Paris headquarters of the United Nations Educational, Scientific and Cultural Organisation (UNESCO) at the end of last month to discuss the prospects for international cooperation on the human genome project. Enthusiasm for the project was there — but not agreement on who should do the work, who should pay for it and, above all, "when to go big", as the European Commission's director-general of science, research and development, Paolo Fasella, put it.

To the surprise and relief of many delegates, James D. Watson, director of the US National Institutes of Health Center for Genome Research, said that an all-out drive for automatic end-to-end sequencing of the human genome would not begin for five years. Eric Barnard, of the British Medical Research Council Molecular Neurobiology Unit at Cambridge, suggested that Europe could wait longer to join the effort to sequence the genome from end to end. But Norton Zinder, of Rockefeller University, expects a substantial contribution to the work from countries other than the United States, which says it will "do and pay for no more than half the work".

Both A. A. Bayev of the Soviet Academy of Sciences and Kenichi Matsubara, Japanese vice-president of the Human Genome Organisation (HUGO), left little doubt for US scientists that, if they wanted to 'go big' immediately, they would probably be alone. According to Bayev, the Soviet Union is handicapped by its lack of hardware. "I hope for a not brilliant but not bad future", he said, "We do not have new generation computers, but we have very very good mathematicians."

Japanese scientists want to support HUGO, the international body set up to coordinate human genome research, but need more time before they can win financial support from the government — despite the threats made by Watson in January that Japanese scientists could not expect to share US data if their government did not make a contribution in keeping with its economy.

Watson has since softened his position. He now suggests that scientists should be entitled to sit on their data. "I hope other countries will join in", he said. "The costs will come down. If we did sequence completely, which we won't, we'll have the fun of looking at the sequence before we release it." How long might participants keep the data to themselves? Six months was the estimate made by Zinder but there was a feeling that this was just another guess.

Peter Coles

Congress cracks down

Washington

A GROWING number of break-ins, fire bombings and threats by animal-rights extremists is fuelling movement in Congress towards a law that would make crimes against animal-research laboratories a federal offence.

Claiming that state and local police forces are not competent to prevent animal rights attacks on research facilities, several members of Congress have recently introduced bills that would allow the Federal Bureau of Investigation (FBI) and other federal police bodies to investigate crimes by animal activists.

The legislators argue that assaults on animal research laboratories are often planned and executed by national animal-rights organizations, and that participants often cross state lines. They note that at least one extremist organization, the Animal Liberation Front, began in the United Kingdom and continues to have ties there. In the past decade, illegal acts by animal-rights organizations have increased more than three-fold, according to statistics compiled by the National Association for Biomedical Research (NABR). In 1989 alone, NABR reports 19 incidents, ranging from sit-ins to bombings. Seven of the more extreme attacks were attributed to the ALF.

Testifying at a hearing on one of the bills last week, Richard Van Sluyters, from the University of California, Berkeley, who has been targeted by animal-rights groups, characterized the activists as "professional agitators . . . Their interstate conspiracy is beyond the local law-enforcement ability", he said.

"When you tell a local policeman that someone has taken your rats and written on your walls, they don't understand the level of tragedy", added Robert Phalen, from the University of California, Irvine.

One bill, introduced by Senator Howard Heflin (Democrat, Alabama), passed the Senate late last year, and has been introduced in the House by Representative Charles Stenholm (Democrat, Texas). Another House bill, which was the subject of last week's hearing, has been introduced by Henry Waxman (Democrat, California).

All the bills agree on one point — breaking into a federally funded research facility would be a federal felony, and solving the cases would be the responsibility of the FBI, the Justice Department or other federal organizations.

Waxman's bill applies only to laboratories funded by the Public Health Service, which would include all facilities supported by the National Institutes of Health. Heflin and Stenholm's bills apply to any research facility, whether private or publicly funded.

But the three bills will probably have to be modified in some way to gain enough congressional support to pass. Heflin's bill, for example, contains language that would make it a federal offence to photocopy laboratory documents without permission, a clause that has been assailed as unduly harsh. "It's important that we don't make it impossible to be a legitimate

ANIMAL RIGHTS: LITIGATION

Activist group under fire

Washington

PEOPLE for the Ethical Treatment of Animals (PETA), an international animal rights group, last week said it would sue for libel over an article in the February issue of *Washingtonian* magazine that alleges that the organization is financially corrupt and that some of PETA's best-known publicity photographs were staged.

The 16-page article, entitled "Beyond cruelty", cites anonymous sources who claim to be former PETA financial officers saying that in 1986 PETA used more than 20 separate bank accounts to hide \$1.2 million in donations that had never been reported as income. The sources also allege the "illegal" use of nearly \$62,000 of PETA funds to pay for the defence and fine of an Animal Liberation Front member who has been convicted of participating in a 1986 University of Arizona break-in.

Katie McCabe, the article's author, also alleges that one of the photographs that PETA has used to symbolize vivisection in its promotional posters was staged. The photograph of a monkey uncomfortably tied to a steel restraining apparatus was taken by PETA founder Alex Pacheco in 1981 in the Silver Spring, Maryland, laboratory of Edward Taub. McCabe alleges that Pacheco, who posed as an interested volunteer to gain access to the laboratory, and a PETA associate tied the animal to the apparatus themselves in a late-night photo session.

Pacheco categorically denies the allegation that there was anything "staged" about the photograph. He says the monkey was tied to the "chairing" apparatus by three workers at the laboratory as was customary. He had been assigned to observe the animal before it was to be experimented on, during which time he took the photographs, he says, but "in no way, shape or form did I modify the apparatus or the animal's position". Pacheco notes that during legal proceedings against Taub the authenticity of the pictures was never challenged. He also points out that the pictures were taken in the afternoon. *Washingtonian* editors say an "editing error" incorrectly place the

whistle-blower", one aide says.

The White House has already stated its opposition to Waxman's bill on legal grounds, and would no doubt oppose the other bills if asked. But in a Democratic controlled Congress, administration opposition is by no means the kiss of death.

Indeed, supporters of the bills believe that some sort of legislation in support of harsher penalties for animal-rights break-ins is likely to emerge this year.

G. Christopher Anderson

event at night.

The *Washingtonian* article has received wide praise from many in the scientific community, and drawn equally strong criticism from animal activists. Last week, James Mason, the assistant secretary of the Department of Health and Human Services, the Federal agency that runs NIH, submitted the article in support of his testimony at a congressional hearing on animal activist legislation (see accompanying article). The article, he said, illustrated the "serious problem" of animal rights extremism. "Not only are the extremists threatening the right of research scientists to work without intimidation but also, and more importantly, they are endangering the future of our nation's health", Mason says.

But activists and their supporters claim that the article errs on many key points, and paints the animal-rights community in an unfair light. "This is clearly a vicious article intended to depreciate the animal rights movement," says Philip Hirschkop, an attorney who says he will file a libel suit this week for PETA. The suit will seek damages from both malicious libel (the allegations that Pacheco had faked the photographs) and business libel (the allegations of PETA's illegal financial practices) Hirschkop says.

Although PETA officials concede that they have several bank accounts, they say that Federal deposit insurance only covers up to \$100,000 per account, thereby forcing them to spread their funds over many accounts. "But all contributions were accounted for and reported to the IRS [Internal Revenue Service]", Pacheco says. G. Christopher Anderson

Eastern Europe: A special issue

A SURVEY of Science in Eastern Europe, compiled by Alun Anderson, Stephen Dickman and John Maddox, will be published in *Nature* on 12 April, in place of the planned survey of Science in France originally planned for that date. The French survey will appear on 12 July.

"Valuable data" claim

Washington

CONTROVERSY over the fate of the seven remaining "Silver Spring" monkeys looks set to intensify following claims that an experiment on one of the monkeys has yielded surprising and important results. Animal rights activists have long argued that nothing useful could be learnt from research on the severely disabled monkeys and that they should be transferred to an animal sanctuary. But the group of researchers who carried out the experiment last month say their data are very valuable and mean that the animals have not suffered prolonged disability to no good purpose.

The group of crab-eating macaque monkeys have been at the centre of a battle between animal rights activists and the National Institutes of Health (NIH) since 1981 when research being conducted by Edward Taub at the Silver-Spring based Institute for Behavioural Research was halted. Taub's experimental monkeys had been deafferented, so that they had lost sensation in one or both forearms, as part of an experiment to assess the effect of training on rehabilitation.

This year, on 14 January, a last-minute court decision (see *Nature* 343, 297; 25 January 1990) finally allowed a group of researchers from the National Institutes of Mental Health and several universities to carry out an experiment on one of the monkeys, named Billy, that had become very sick.

The researchers stress that the experiment was designed in accordance with an agreement reached between the Secretary of the Department of Health and Human Services and members of Congress that the animals would not be used for invasive procedures from which they would be allowed to recover. Euthanasia of Billy had already been recommended by veterinarians as his condition deteriorated. Experiments were conducted under anaesthesia and Billy was then allowed to die.

Preliminary accounts of the experiment given by Timothy Pons of the National Institutes of Mental Health in Bethesda, Maryland, reveal that an unprecedented degree of reorganization of Billy's sensory cortex had occurred. An 8–10-millimetre wide area of cortex that would normally receive input from the hand was found to have completely filled in with input from the face.

The great degree of change in cortical organization is presumably a result of the very long period for which Billy was without sensory input from his forearms. Ironically, Billy had been kept alive with his disability for a decade only as a result of the political and legal problems that mired NIH as it struggled with animal rights activists. Under the original plan for

the experiment, Billy would have been euthanatized many years ago and the experiment would not have been possible.

The researchers disagree utterly with claims by animal rights organizations that their experiment was concocted to help justify NIH's long-standing refusal to hand over the monkeys to an animal sanctuary. They point out that they approached NIH at their own initiative in 1987. They say that they had realized that the new knowledge then emerging about cortical reorganization made the animals particularly valuable in a way that was largely unanticipated when Taub planned his own experiment eight years earlier.

Opponents of the experimental use of the monkeys remain unconvinced. Neal Barnard, president of the Physicians Committee for Responsible Medicine (PCRM) which last month unsuccessfully called for an enquiry into the experiment by the DHSS Office of Scientific Integrity Review, says the result is predictable and of no great significance.

Barnard says that during earlier court battles he told a judge that "NIH desperately needs data and although we all know that brain reorganization has long been known to occur they will show 'wow we've got brain reorganization over one square centimetre isn't that phenomenally amazing' . . . The only people who are enthusiastic about the results are other people doing these brain-mapping studies," says Barnard, "Science has become secondary to public relations and politics".

Not so argues Pons. "Anyone who does these kinds of mapping studies and would know what to expect would be surprised, if not shocked, by the result. In fact, there has never been a report of complete reorganization, to my knowledge, after deafferentation, especially in an adult. It has never been predicted."

David Hubel, Nobel laureate and Professor of Neurobiology at Harvard Medical School, describes the result as "clearly important". Earlier work, he notes, has shown a smaller degree of cortical reorganization which was rather close to the resolution of the method. Here, several hundred recordings indicate an 8mm square piece of cortex has filled in with input from the monkey's face. He adds that it would have been "almost criminal to let the monkey simply die, when you could conduct an experiment without the least suffering on the part of the animal."

Alex Pacheco of PETA says he wants to see the work in print and hear the opinions of independent neuroscientists before he makes a judgement. "NIH has too much of a vested interest", he says, "they have no choice but to say they are learning great things".

Alun Anderson

UN fragments over treaty

Washington

In protest at the slow pace of United Nations (UN) global warming negotiations, a splinter group of nine countries last week said they would soon begin work on their own treaty to limit greenhouse gases, with the hope that the United States and other nations will eventually join in.

The group hopes to begin treaty negotiations in late summer, after the report from last week's UN climate change meeting in Washington is completed. "We don't need 'final scientific proof'. We think there is already enough evidence of a global warming", says Heintz Schreiber, head of the Austrian delegation to the UN Intergovernmental Panel on Climate Change (IPCC).

The Austria-led initiative already has the support of Denmark, France, West Germany, Italy, the Netherlands, Norway, Sweden and Switzerland. Further participants are expected, but the sponsors acknowledge that they do not know exactly how many, or which, countries will be necessary to achieve the "critical mass" that could result in worldwide endorsement. "the groups hopes to have a treaty before the end of 1991.

In a statement, the group said that it is "concerned that any delay in action may endanger the future of further generations of mankind". It hopes that, with a working treaty in place, major industrialized nations such as the United States and Japan — as well as many developing countries — will be forced by public opinion to agree to greenhouse gas limits.

Although the group has set no specific targets for emissions, they intend to use the developing consensus that emerged from last week's IPCC meeting as a foundation. The IPCC scientific panel is preparing a draft report that is expected to predict temperature rises of between 1.5 and 4.5 °C by 2030 and a sea level rise of 20–30 cm, given a doubling of CO₂ levels.

A report released by Greenpeace at the meeting concludes that, with substantial emission-control measures, global temperatures could eventually stabilize at between 1.1 °C and 3.1 °C above the pre-industrial level. Among the controls that would be necessary are halting chlorofluorocarbon production by 1995, reducing energy-related CO₂ to 30 per cent of the present value by 2020, and extensive reforestation, the 80-page peer-reviewed report finds. The report's author, Mick Kelly of the Climate Research Unit of the University of East Anglia, says such restrictions "would require a major shift in energy policy, but [are] not beyond the bounds of possibility".

G. Christopher Anderson

Trouble building at Harima

Tokyo

PLANS to build the world's largest synchrotron in a new 'science city' in western Japan ran into their first problems last week when a group of construction companies were accused of rigging their bids for ground preparations for the project.

The 8-GeV Synchrotron Radiation Facility is being built by the Science and Technology Agency (STA) in the Harima Science Garden City in Hyogo Prefecture west of Osaka. The garden city will form the core of the Nishi-Harima 'Technopolis', one of several technopolises being established by the Ministry of International Trade and Industry (MITI) and local governments to assist the economic revival of Japan's remoter regions (see *Nature* 329, 478; 1988).

Preparation of the grounds for the 2,000-hectare garden city will be paid for by the Hyogo Prefectural government and is expected to cost ¥120,000 million (\$850 million), according to Tohru Amano, director of STA's office for construction of the synchrotron. The synchrotron itself will cost STA about ¥100,000 million (\$715 million), considerably more than

...and if there's a Nobel Prize,
we split the money three ways...



the cost of Japan's present biggest accelerator TRISTAN at the High Energy Physics Laboratory (KEK) of the Ministry of Education, Science and Culture (MESC) in Tsukuba.

Last week's bids were for ground preparations for the 141-hectare site for the synchrotron. The site was divided into two lots for bidding. But on the day before tenders were submitted, an anonymous letter identifying the successful bidders was sent to the Hyogo prefectural government and the mass media. One bid of ¥4,270 million was submitted by Ohbayashi Corporation, Konoike Corporation, Arai-Gumi, Ltd, Konoike Construction Co. and a local company. The other successful bid of ¥4,440 million was made by Nishimatsu Construction Co., Kajima Corporation, Yoshida-Gumi, Ltd, and another local company. These bids were about ¥100 million less than those submitted by nine other groups.

A spokesman for Ohbayashi Corporation was quoted on national television as saying that "there is nothing to indicate that the company was involved in irregular procedures". But Japan's construction industry is notorious for its practice of 'dango', a process whereby big construction companies get together before sub-

mitting tenders and make a decision about who will win the bid. This is the practice that has been attacked by the United States for effectively shutting foreign companies out of Japan's gigantic construction market.

The Hyogo prefectural government has said that the matter will be thoroughly investigated and Amano is confident that the local government will resolve the situation.

David Swinbanks

Hundreds keen to get SOR

Tokyo

JAPANESE scientists are already lining up in their hundreds to gain access to the 8-GeV Synchrotron Radiation Facility even though the synchrotron will not undergo its first trials until 1995 and is not expected to begin full operations until three years after that.

When completed, the Synchrotron Radiation Facility (SOR) will be more powerful than both the 7-GeV Advanced Photon Source (APS) at the Argonne National Laboratory in the United States, scheduled to be built by 1995, and the 6-GeV European Synchrotron Radiation Facility (ESRF) at Grenoble, beginning operation in 1993.

According to Amano, about 400 researchers from universities, government and industry have already joined working groups that are now planning use of the synchrotron.

Construction of the facility is being directed by STA's Institute of Physical and Chemical Research (RIKEN) and the Japan Atomic Energy Research Institute (JAERI), another affiliate of STA. But once completed, the synchrotron will be taken over by a new organization and will be open to use by researchers throughout Japan. Amano says that 'nearby' foreign countries may get access to one or two of the synchrotron's 50 beam lines. Preliminary discussions have been held with researchers from Korea, Taiwan and Australia.

The beams from the synchrotron will be able to resolve features down to the size of an atom, or about one ångström. The facility will be particularly useful for determining the structure of proteins. Japan's present most powerful synchrotron, the Photon Factory at KEK in Tsukuba, a 2.5-GeV facility, takes two to three days to determine the structure of a protein but the same job will take only take two to three minutes at the new facility, according to Amano.

The 8-GeV Synchrotron Radiation Facility is too powerful to use for development of sub-micrometre lithography of VLSI (very-large-scale integration) circuits, a common application for present-day synchrotrons. But Amano says there is plenty of space next to its 1-GeV linear accelerator for smaller synchrotrons that

could be operated by industry, and some companies have shown interest in attaching additional facilities to the synchrotron's linear accelerator. But Sumitomo Heavy Industries and Hitachi have already developed self-contained mini synchrotrons for VLSI development and it is uncertain what role the Synchrotron Radiation Facility will play when it comes into use in eight years' time.

David Swinbanks

PRIZES

Japan prize awarded

Tokyo

ONE of Japan's most lucrative science prizes, the Japan Prize, has been awarded to Marvin Lee Minsky of MIT in the United States who is credited with developing and popularizing the concept of artificial intelligence. A group of three geologists who helped develop the plate tectonic theory, Dan Mackenzie of Cambridge University in the United Kingdom, Xavier Le Pichon of Ecole Normale Supérieure in France and

Minsky of MIT (right) gains a Japan Prize for work on artificial intelligence. Mackenzie (below left), Le Pichon (below right) and Morgan are recognized for work on plate tectonics.



William Jason Morgan of Princeton University, share the other of the two awards announced last week.

The prize is administered by the Science and Technology Foundation of Japan and each award is worth ¥50 million (\$345,000). The prizes will be presented in Tokyo on 17 April.

David Swinbanks

PASTEUR INQUIRY

Cancer risk indicated

Paris

ACCORDING to the preliminary results of an inquiry directed by Professor Jean Bernard, personnel at the Pasteur Institute in Paris could be more susceptible to rare forms of cancer than the French population as a whole. Although the 12-member committee of inquiry says that further investigation is necessary, the results of the inquiry will strengthen concern about the seven Pasteur researchers who developed rare forms of cancer within a few years of each other. Two of them have since died (see *Nature* 338, 607; 1989).

The inquiry was set up in 1987 with the aim of tracking down as many as possible of the technical and research staff who had worked at the Pasteur Institute during the same period (1971 to the end of 1986) as the sick researchers. These researchers had worked in laboratories on the same floor and some of them had handled recombinant DNA products. Epidemiologists from Unit 170 of INSERM (the National Medical and Health Research Institutes) finally traced 3,765 individuals fitting the criteria.

Only mortality statistics have been analysed so far in the study. Out of the sample, 143 deaths were recorded, compared with 246 expected for the same age range in the French population as a whole. This lower death rate, says Annie Sasco, one of the epidemiologists from the Lyons laboratory which carried out the inquiry, is probably linked to the sample's higher-than average socioeconomic status. Of these 143 deaths, 51 were from cancers and 41 from cardiovascular disease. Both of these proportions are somewhat less than would be expected for the population as a whole.

Deaths from commoner forms of cancer were less frequent than expected, but those from cancers affecting the bone, the pancreas and the brain were higher. Two deaths from bone cancer were found, compared with an expected number of 0.58, while six of the sample died from cancer of the pancreas, compared with an expected 2.67.

The sample data do not permit conclusions about risks at the Pasteur, said Sylvaine Cordier, an epidemiologist who was also part of the inquiry committee. The study needs to be followed up by research into morbidity, she said.

Other people may have developed cancer and been cured, or may be developing illnesses that have not yet proved fatal. Further studies taking into account environmental factors, such as the kinds of products handled or exposure to radiation, are now necessary. **Peter Coles**

ISRAEL

Soviet influx opens new era

Jerusalem

SCIENCE and technology are certain to be profoundly affected by the massive influx of Soviet Jews which is already beginning to change the face of Israel. As many as one million émigrés may arrive before the end of the decade, increasing the Israeli population by a quarter and tripling the size of the scientific community.

Baruch Eyal, head of human resources and international research institutes at the Science and Technology Ministry, says that the official forecast figures are "quite astounding; for every 100,000 immigrants, 42,000 are professionals, including a particularly high percentage of engineers and about 2,000 scientists, 1,500 of whom are in natural and exact sciences, with the remainder in social sciences. This is five times higher than the concentration of scientists in Israel and the West generally." In January alone, 6,170 Soviet citizens landed in Israel, and the proportion of scientists among them reflects Eyal's forecast figures.

The Science and Technology Ministry has been formulating plans to absorb these professionals and welcomes proposals from scientists still waiting to leave the Soviet Union. Ministry officials claim to be making special efforts to ensure that the émigrés stay in Israel rather than move on to the higher salaries in the United States and Canada. But scientists in industry accuse the government of "laziness" and argue that it is still not doing enough to prepare for the great influx.

Eyal responds by saying that "so far, the ministry has earmarked an initial \$3 million for the employment of Soviet scientists, with a focus on the fields of biotechnology, materials, computer science and artificial intelligence, electro-optics and lasers, neurobiology, superconductivity, and environmental sciences". A Soviet scientist who finds work, whether in a university or in industry, will be able to keep that government salary rather than strain the budget of the institution in question.

The Israeli Academy of Sciences has also announced the allocation of \$6.3 million, through a private fund, for 15 senior Soviet scientists. The money will be used to establish chairs in their fields of expertise and for housing loans and five-year research grants.

A second ministry plan is to strengthen existing infrastructure. Israel's seven universities are short of funds and do not have the means to increase their staff but "no one wants to waste the resources", says Eyal, "especially in the case of 'exceptional' scientists and those with specific know-how, such as biotechnologists." There are plans to set up a biotechnology institute within the Institute of Applied Sciences at

Beersheba, and possibly an institute for opto-electronics.

Software engineers — which in the Soviet case means mathematicians — will be much in demand, given the potential for creating exportable items without massive investment in plant. Efforts will also be made to encourage scientists to move into school teaching.

To facilitate implementation of the ministry's plans, a new computer network will link the ministry with universities and industry and representatives in Moscow.

The influx of immigrants is likely to increase competition for jobs. Eyal stresses that it is "important that a gap does not develop between immigrants and demobilized soldiers, new graduates and other Israelis competing for the same jobs". But Shmuel Adler, who heads the Absorption Ministry's Centre for Absorption in Science, warns that "Israelis have to learn that in the question of manpower, the best one will win." Since the centre was set up in 1974, Adler says that it has helped 4,500 scientists start work, half of them from the Soviet Union.

One enterprising example of 'scientific absorption' was initiated by Herman Branover, professor of mechanical engineering at Ben-Gurion University of the Negev, in the southern city of Beersheba. His Jerusalem-based Shamir Advanced Technologies Engineering Centre (SATEC) employs about 60 people, of whom 35 are new immigrants. Its principal products are digital electric power meters, low-technology heating elements and hydro-metallurgic techniques for extracting precious metals.

SATEC's scientists are all of Soviet origin — including Branover who arrived in 1972 — while the company's marketing and finance is handled by American- and British-born Israelis who immigrated with skills that native Israelis could not provide.

Branover says that not enough people in government or industry appreciate the scale of the new migration. "We are expecting a million people; and 25 per cent have engineering, medical or technical degrees; 2–3 per cent have at least doctoral qualifications". It is possible, Branover believes, for Israel to take a "quantum leap in high-technology fields like that seen in Japan or Korea". As for the immigrant scientists themselves, Branover says, "they are refugees. They are ready to make concessions." Government officials agree, adding that this is not the generation of 'refuseniks' who were fighting for freedom of expression and religion. In Adler's words, "now we are talking about an exodus, not an ideological *aliya* [immigration to Israel]".

Lisa Perlman

Merger plan dropped

London

JOHN MacGregor, the UK Secretary of State for Education and Science, has decided not to set up a working group to consider the possibility of a merger between the Agricultural and Food Research Council (AFRC) and the Natural Environment Research Council. The Advisory Board for the Research Councils (ABRC) recommended to MacGregor in a letter on 1 December that a small group under ABRC should examine the idea of a closer association between the two research councils, as a way of eliminating turf-battles over responsibilities for environmental science, and especially land use research.

MacGregor's reply notes that the streamlined ABRC, which will start work on 1 April (see *Nature* 343, 294; 1990), has been set up to improve cooperation among all five research councils within the existing research council structure, a matter to which he expects the new board "to give urgent attention". P.A.

Alternative on loans

London

THE UK Committee of Vice-Chancellors and Principals (CVCP) this week unveiled an alternative to the government's student 'top-up loans' scheme. The CVCP proposes that students should receive an adequate grant, not means-tested against parental income, but should pay a 'graduate tax', once their income is higher than the national average.

The government's scheme aims to keep the present maintenance grant with a contribution from parents, dependent on income. But the grant would be frozen at its current level, and supplemented progressively with centrally provided loans.

The CVCP says that it is willing to run its scheme, which, it claims, would be cheaper to run than the government's proposals. Universities would provide grants, and be reimbursed by the Department of Education and Science, either directly or through the Universities Funding Council.

A spokesman for the CVCP says that, in view of the difficulties the government is having in setting up its student loans company, following the major banks' refusal to become involved, the time is right to consider alternatives. But the CVCP's scheme may be unacceptable to the government, which has stated previously its opposition to any kind of graduate tax. P.A.

Old flames

A FIRE destroying the "fine buildings" of the University of Toronto was in the news this week, 100 years ago. Ten years ago, President Carter was easing back on the US commitment on fast-breeder reactors. See the Then and Now page, facing the inside back cover. □

Marshall pulls no punches

London

BLAME for the abandonment of the British pressurized water reactor (PWR) programme rests with the British government, according to Lord Marshall, former chairman of the Central Electricity Generating Board (CEGB). He was giving evidence to the House of Commons Select Committee on Energy last week.

Marshall said the particular formula chosen for the financial structure of the privatized electricity industry had led to escalating estimates of the cost of nuclear power and the subsequent scrapping of plans to build further PWR generating plant (see *Nature* 342, 213; 1989).

The sale to public shareholders will begin in April, when twelve distribution companies will be sold. Privatization of the National Grid and of nationalized generating plant (in three parcels, one in Scotland and two in England and Wales) will follow next year.

Marshall contrasted the British plan with arrangements in the United States and Japan, where generating companies hold a franchise for a geographical area in return for an 'obligation to supply'. The crucial question, he said, was British government's original plan to place the obligation to supply with the distribution



Marshall: No looking back . . .

companies, holding regional franchises, but with no long-term obligations to the generators.

At that stage, Marshall said, he had thought privatization of the nuclear industry was "difficult, but not impossible". But when the government later chose to allow free competition between distributors, removing the obligation to supply, successful privatization of nuclear plants became impossible.

The government "should have been aware . . . that they were going to make life very difficult for nuclear power", Marshall said. The final form of privatization meant that the distributors would not have signed long-term contracts for nuclear generated electricity with National Power (the company to have taken over Britain's nuclear generating capacity). But without long-term contracts or government guarantees, the large

initial capital costs of PWR reactors would have to be recouped over a much shorter period than their operating life, inflating short-term costs.

The need to make a competitive rate of return on invested capital in the private sector (about 10 per cent a year, rather than the 5 per cent demanded by government during most of the 1980s), Marshall said, had also increased cost of nuclear power.

But Marshall admitted that, even if the electricity industry had been privatized as he had wanted, PWRs might still have been more expensive in the short term than gas turbine power stations. "In my view, we would have built nuclear plants and gas turbines simultaneously", he said.

Asked if he could have been more helpful to the government during the build-up to privatization, Marshall regretted he had been "too amenable".

Marshall also attacked the management of Britain's nuclear programme by successive governments, in particular the decision to develop advanced gas-cooled reactors (AGRs).

But Marshall, himself, was criticized during questioning for failing to obtain "substantive discussions" on the privatization of nuclear plants with the Department of Energy, during the early stages of privatization. Marshall answered that Department officials had been busy with the non-nuclear side of privatization during 1988 and the early part of 1989. Discussions were also delayed until CEGB had received figures from British Nuclear Fuels for fuel reprocessing costs.

Increases in these reprocessing costs were blamed by Marshall for the spiralling estimates of the cost of decommissioning Britain's ageing Magnox reactors, responding to suggestions that only privatization had forced CEGB to reveal the true costs of decommissioning.

Marshall explained that a shut-down Magnox reactor contained the equivalent of 8 year's fuel. British Nuclear Fuels traditionally had a 'cost-plus' contract with CEGB, where any increased costs, for example due to environmental safety improvements, were simply charged to CEGB. As privatization approached, British Nuclear Fuels was forced to give fixed-price quotes for reprocessing, which were set at a high level to cover future cost increases.

Despite this, Marshall said he had been surprised at the government's initial decision in July 1989 to remove Magnox stations from privatization (before the decision to retain all nuclear plants in public ownership). The most natural decision, he believed, would have been to retain the fuel cycle in the public sector and privatize the reactors. **Peter Aldhous**

SOUTH AFRICA

Separate education persists

Cape Town

THE South African government has released Nelson Mandela, ended the ban on the African National Congress and allowed the return of political exiles, but it is still not prepared to deracialize its education system.

The Minister of (African) Education and Training, Stoffel van der Merwe, told a briefing for journalists last week that this would create "tremendous chaos," and would have a negative effect on both political stability and the economic system. Once the different education departments had reached a point approaching equality, he indicated, the separation of the various departments could be re-examined.

It is difficult to see just how the present government envisages this happening during its lifespan. In 1986 it embarked on a ten-year programme to eliminate disparities in per capita expenditure on education between racial groups. But last year it acknowledged that this goal was economically unfeasible: current per capita spending on white schoolchildren is approximately five times that on African schoolchildren, and whites comprise only 10 per cent of the 9 million pupils in the country.

These discrepancies are no longer reflected so much by differences in standards between examining boards, which have narrowed in recent years, but by pupil:teacher ratios, teachers' qualifications and subjects offered. For example, only 30 per cent of African pupils study mathematics at any level, and in 1988 only 764 passed the subject in the higher grade, compared to approximately 12,000 whites.

Now van der Merwe says that money

released as a result of cuts in defence spending will be directed towards black education, and he hopes that this will be supplemented in future by funds likely to become available as a result of the possible relaxation of sanctions.

Although the figures will not be announced until the budget next month, it is suspected that these funds will be required simply to meet the demand for increased facilities: an estimated one million African children have no schools to go to and 300 new schools and 8,000 additional teachers are needed each year to meet demand from the 250,000 African pupils entering the system. In contrast, the white school-going population is declining by 9,000 per annum, and there are currently 308,000 places open at white schools (26 per cent of places available). In addition, the government has closed 196 white schools over the past decade.

But they continue to refuse state schools the option of becoming non-racial, despite the fact that several leading ones have expressed a desire to do so.

In Port Elizabeth last week, 100,000 people marched on the offices of the Department of Education and Training to demand the opening of schools to all races. In this eastern Cape city, 3,000 African pupils were unable to be accommodated when the school term commenced last month. In response to this, local communities mustered up to 40 unemployed teachers in the city who were prepared to teach them at existing schools by adding an additional shift to the day's programme. The Department of Education and Training responded that it simply did not have the funds to pay their salaries.

Michael Cherry

SATELLITES

Successful Japanese launch

Tokyo

JAPAN's National Space Development Agency (NASDA) last week successfully launched a marine observation satellite (MOS-1b). The satellite will provide high-resolution images of ocean currents, temperature and sediment distribution in the sea surface and of vegetation, snow cover and geological features on land. It will also collect information on cloud and



Japan's moon launch of a few weeks back.

water vapour distribution in the atmosphere. The Sun-synchronous orbit will allow the satellite to scan the whole surface of the Earth every 17 days.

The H-1 launch vehicle also deployed a small satellite for use by amateur radio operators and a Deployable Boom and Umbrella satellite which will repeatedly extend and retract a 95-cm diameter umbrella to test out techniques for braking and altering the orbit of satellites.

David Swinbanks

INDIA

Turning to turtles to clean Ganga

Bangalore

CARNIVOROUS turtles are to be released into the River Ganga in an effort to help reduce the effects of the age-old practice of throwing half-burnt human bodies into the river. A Rs2,920-million project to clean up the heavily polluted Ganga is already underway but has had difficulty finding solutions to the problem of floating corpses in the river.

"Since the problem of getting rid of dead flesh defied technical solutions, raising turtles seemed a safer biological option", says a spokesman for the clean-up project. About 7,000 meat-eating turtles are being reared in a special farm and will be released into the river. The plan is considered essential because pious Hindus who regard the river as divine are thought unlikely to make much use of the 32 electric crematoriums being set up along the banks of the river.

Radhakrishna Rao

MIT PRESIDENT

Leading candidate drops out

Boston

THE search for a new president of the Massachusetts Institute of Technology narrowed recently when MIT provost John M. Deutch, regarded by many as the leading candidate, announced that he will resign in June and will not be MIT's next president. Insiders believe his action was precipitated by indications from the search committee that he would not be selected, and a similar explanation has been offered for his decision earlier last month to withdraw from consideration for the presidency of Johns Hopkins University.

Because MIT and Johns Hopkins are among the biggest recipients of Pentagon funds of any universities in the United States, it had been assumed that the school search committees would view

Deutch's well-known involvement with the US Defense Department as an asset. Deutch has, for nearly twenty years, served on the Defense Science Board, a group which advises the secretary of defense. But recent student and faculty criticism of Deutch's close ties with the military and with industry may have hurt his chances.

Some Johns Hopkins faculty reportedly raised objections to Deutch's candidacy in a time of "lessening world tensions and diminishing defense budgets." Earlier this year a student newspaper on the MIT campus also raised controversial questions about the extent of Deutch's corporate ties, alleging that corporate directorships and outside consulting work account for more than half of his income.

Seth Shulman

Agent Orange controversy

SIR—The report “‘Not guilty’ verdict challenged” (*Nature* 342, 217; 1989) on the *Evatt Revisited* conference¹ correctly identified one of the reasons for continuing controversy about the Royal Commission into the Use and Effects of Chemical Agents on Australian Personnel in Vietnam: the use without acknowledgement in its final report, which concluded “Agent Orange: Not Guilty”², of large tracts of Monsanto’s submission³⁻⁸. This has, however, been known for some time³⁻⁷; the main point of the *Evatt Revisited* conference was a reappraisal of scientific evidence presented to the commission, and published since. In the introductory paper, my colleagues and I concluded that on balance the evidence now suggests that pesticide exposure increases the risks of some cancers and birth defects, but that there is uncertainty about this conclusion that, at least in the case of the Vietnam veterans, is unlikely to be resolved by further research⁸. This is consistent with reports since the *Evatt* Commission by the International Agency for Research on Cancer and the USAF School of Aerospace Medicine, as well as several research papers⁹⁻¹⁶, but other authors still defend the commission’s unqualified exoneration of Agent Orange, dioxins and other chemicals used in the war¹⁷⁻²¹.

Which of these studies should be accepted, and what level of proof to apply, is a social choice on which scientists can only advise; there is no statistical algorithm that will convert a controversial and confused situation into a simple objective truth that absolves society from making choices that involve value judgments. The traditional scientific stance of scepticism and critical rigour guards against errors of credulity, and in this case happens to favour chemical companies and governments against the veterans; but it is possible to be in error because of excessive scepticism as much as credulity, and in this case to deny the veterans just treatment.

In *The Politics of Agent Orange*²², McCulloch argued that epidemiology “cannot resolve the problem of public responsibility for the suffering of the veterans”, and speculated prophetically: “Perhaps the specialists will reveal eventually that it is not possible to prove or disprove that the veterans are ill because of chemical exposure which occurred in the RVN [Republic of Vietnam]”. It is obvious that social decisions based on conflicting scientific evidence and opinion must be made under uncertainty (many philosophers would argue that all decisions are inevitably of this type), and involve issues of trust, justice and equity as well as theoretical knowledge and technical competence. To me the critical ethical issue is the question raised at *Evatt*

Revisited by Professor Axelsson as well as myself: “who gets the benefit of the doubt?” The organizers of *Evatt Revisited* have presented social justice arguments that the veterans should be given the benefit of the doubt in this case⁸.

A. J. D. BELLETT

*Division of Cell Biology,
John Curtin School of Medical Research,
The Australian National University,
Canberra, ACT,
Australia*

1. Steele, E.J., Bellett, A.J.D., McCullagh, P.J. & Selinger, B. (eds). 1989. *Evatt Revisited*: interpretation of scientific evidence. Proceedings of a conference which re-examined the findings of the Royal Commission on the use and effects of chemical agents on Australian personnel in Vietnam (Centre for Human Aspects of Science and Technology, University of Sydney, 1989).
2. Royal Commission on the use and effects of chemical agents on Australian personnel in Vietnam. Final Report (The Hon. Mr. Justice Phillip Evatt, Commissioner). (Australian Government Printing Office, Canberra, 1985).
3. Axelsson, O. & Hardell, L. *Med. J. Aust.* 144, 612 (1986).
4. Pollak, J.K. *Med. J. Aust.* 144, 612 (1986).
5. Martin, B. *Aust. Soc.* 5 (11), 25 (1986).
6. Hardell, L. & Axelsson, O. *Med. J. Aust.* 145, 298 (1987).
7. Humphrey, G.F. *Aust. Soc.* 6 (3), 46 (1987).
8. Bellett, A.J.D., McCullagh, P.J., Selinger, B. & Steele, E.J. in *Evatt Revisited* (1989).
9. International Agency for Research on Cancer. *Working Group on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Some Halogenated Hydrocarbons and Pesticide Exposures* (Lyon, 1986).
10. Albenese, R.A. *United States Air Force Personnel and Exposure to Agent Orange. Interim report for period March 1984 — February 1988* (USAF School of Aerospace Medicine, Division, Brooks Air Force Base, Texas 1988).
11. Field, B. & Kerr, C. *J. med. Genet.* 25, 819 (1988).
12. Axelsson, O., Persson, B. & Hardell, L. in *Evatt Revisited* (1989).
13. Woods, J.S., Polissar, L., Severson, R.K., Heuser, L.S. & Kulander, B.G. *J. natn. Cancer. Inst.* 78, 899 (1987).
14. Hoar, S.K. *et al. J. Am. med. Ass.* 256, 1141 (1986).
15. Buckley, J.D. *et al. Cancer Res.* 49, 4030 (1989).
16. Blakley, P.M., Kim, J.S. & Firneisz, G.D. *Teratology* 39, 237 (1989).
17. Armstrong, B.K. *Med. J. Aust.* 144, 284 (1986).
18. Young, A.L. & Reggiani, G.M. (eds) *Agent Orange and its Associated Dioxin: Assessment of a Controversy* (Elsevier, Amsterdam, 1988).
19. Hay, A. *Nature* 340, 353 (1989).
20. Donovan, J.W. in *Evatt Revisited* (1989).
21. MacPhee, D. & Hall, W. in *Evatt Revisited* (1989).
22. McCulloch, J. *The Politics of Agent Orange. The Australian Experience*. (Heinemann, Richmond, Victoria, 1984).

Reprint solidarity

SIR—I agree with Alexander C. Leung (*Nature* 339, 574; 1989) that “reprints are costly to authors and should not be requested unless there is a really good reason for it”. However, for most of us working in Poland (and I believe in most Eastern European countries) reprints are the only available source of scientific information. As a neuropathologist and electron microscopist, I have no access to any neuropathological or electron microscopical journal, let alone books.

I cannot read an article before requesting a reprint, as I know the title of the article only through *Current Contents* or from the bibliography section of another article. Even sending a request for a reprint is expensive, as in many situations I have to pay for requests from my own

pocket. Even if I have access to journals, I would not be able to make a photocopy as such a service is almost non-existent in scientific libraries in Poland. It could be said that working in such conditions is my choice, but every time I obtain a reprint I have a feeling of understanding and international ‘solidarity’ among scientists.

PAWEŁ P. LIBERSKI

*Medical Academy Lodz,
Department of Oncology,
Electron Microscopy Laboratory,
Gagarina 4,
Lodz, Poland*

CF screening

SIR—We disagree with L. P. Ten Kate (*Nature* 342, 131; 1989) who argues that screening of cystic fibrosis (CF) carriers should be delayed until detection of 96 per cent of all mutated genes would be possible. It is certainly true that anxiety of couples with one carrier of the F_{508} deletion would be increased by knowledge of their situation. Nevertheless, we consider that it would be unethical to neglect the possibility of informing about 50 per cent of the couples with a 0.25 risk of having a CF child that an ante-natal diagnosis is available to them. We propose that couples should be offered the possibility of being screened for the F_{508} deletion while being explicitly told that (1) the procedure will fail to detect the risk in about 50 per cent of the cases, and (2) only if both of them were carriers of the mutation would they be informed of a positive result. By doing so, the possibility of helping 50 per cent of couples at risk of having a CF child would not be sacrificed to the psychological comfort of the others.

GILBERT VASSART

PASCAL COCHAUX

MARC ABRAMOWICZ

*Service de Génétique Médicale,
Université Libre de Bruxelles,
Hôpital Erasme,
Bruxelles,
Belgium*

Fine writing?

SIR—In your leading article “Adirondacks awake” (*Nature* 343, 101; 1990), you mention the *New Yorker*’s “devotion to fine writing” and its commissioning of “fine writers”. The first sentence in the third paragraph, a quotation from one of these “fine writers”, reads “A forty-minute hike brings the dog and I to the top of the hill behind my home”. Presumably if the writer had gone unaccompanied, it would have read “A forty-minute hike brings I to the top of the hill behind my home”. Perchance the big apple hath a worm betwixt its rosy cheeks.

BRADFORD D. PEARSON

*7431 Lindsay Road,
Richmond, BC V7C 3M7, Canada*

Problems of development aid

SIR—The world is experiencing more frequent collapses of natural resource systems leading to disasters that require international relief. But many of these disasters are themselves the ecological consequences of previous 'development' aid¹. The reasons are illustrated by a recent paper from the International Institute for Environment and Development² on evaluating success in resource management projects in developing countries. We are not criticizing the paper, but use it to point to a widespread limitation in thinking about development aid, namely lack of attention to space and time scales in assessing the dynamics of ecological systems.

The problem is this: the success of aid projects has been judged (1) by criteria more relevant to the donor country than to the recipient, such as completion of wells and dams; and (2) subjectively by outsiders, on single visits, with conclusions reached by consensus. Skinner² has suggested alternative criteria that include whether (a) resource use is sustainable, (b) local people benefit in their own estimation and (c) projects meet local government goals. But, these criteria also cause, in the long term, rapid population increase, negation of project benefits, degradation of ecosystems and increase of poverty¹.

In the Sahel famine disasters, much of the blame ascribed to drought is misdirected. The 'droughts' were man-made and the ecological basis for the breakdown of the ecosystem has been described^{2,3}. In general, the number of organisms that any ecological system can support over time depends partly upon their spatial dynamics. Systems with spatial migrations can support more consumers than equivalent systems where organisms are locally constrained, such as human agro/pastoral systems. Spatial and temporal variation are essential requirements for the persistence of most natural ecosystems^{4,5} yet it is an almost universal aim of policy planners and managers to prevent fluctuations and to spread things out evenly in space.

In Africa, most of the attention paid to aid has been directed towards regions surrounding the Sahara. Without the resulting permanent water supplies and settlement schemes, we contend that much of the disaster would not have occurred. But that is history. Aid organizations are now preparing to repeat the process around Africa's southern Namib-Kalahari desert. As scientists with experience of these ecological systems, we are apprehensive about what will happen to Namibia, Angola and adjacent regions in Botswana when political change allows the next round of arid region 'development'.

We urge restraint in development of

permanent water supplies, use of fertilizers and so on until a thorough analysis is made of the likely ecological consequences over time and space. An essential question, which should be asked before the project is approved, is: 'Will the proposed development change the temporal/spatial dynamics of the systems? More specifically, will it lead to changes in the distributions of livestock, water (from rainfall) and soil?' If the answer is 'yes', it should be reconsidered, bearing in mind the lessons we have already learned^{1,3,5}.

B. H. WALKER

Division of Wildlife and Ecology, CSIRO,
Box 84, Lyneham, ACT 2602, Australia

A. R. E. SINCLAIR

Department of Zoology,
University of British Columbia,
Vancouver V6T 2A9, Canada

1. Sinclair, A.R.E. & Wells, M.J. in *Food and Natural Resources* (eds Pimentel, D. & Hall, C.) 439–484, (Academic, New York, 1989).
2. Skinner, J.R. Int. Inst. for Envir. Dev. Pap. No. 5, (IUCN, Switzerland, 1989).
3. Sinclair, A.R.E. & Fryxell, J.M. *Can. J. Zool.* **63**, 987–984 (1985).
4. Walker, B.H. in *Conservation for the 21st Century* (Western, D. & Pearl, M. eds) (Oxford University Press, Oxford).
5. Walker, B.H., Emslie, R.H., Owen-Smith, R.N. & Scholes, R.J. *J. appl. Ecol.* **24**, 381–401 (1987).

University funding

SIR—D. T. Elmore has questioned the validity of the negative association between Universities Funding Council (UFC) research grades and the *Financial Times* graduate employment statistics in UK universities (*Nature* **342**, 729; 1989), but his report of my statistical analysis (*Nature* **341**, 562; 1989) is misleading. The negative correlation is present regardless of whether extreme outlying points are excluded from the analysis and the collation of data for colleges of the Universities of London and Wales is reported clearly, being based on the numbers of departments in each college.

The additional points made by Elmore relate to variability in graduate employment prospects and to the inadequacies of the UFC research grading process. But these speculations are not relevant to the interaction between graduate employment and UFC research grades. While I agree with Elmore that both data sets are imperfect, it is difficult to see how their inadequacies would produce any significant relationship between graduate employment records and research status, let alone a negative one.

I am forced to conclude that "in general, poor research gradings are associated with enhanced graduate prospects of gaining a foothold in the world of work". This may not be what the academic community wishes to hear, but, on the basis of the best data available, I contend

that the effect is real.

Elmore suggested that my analysis is "an open invitation to government . . . to downgrade some departments or even universities unjustifiably to a teaching role only". I am not in the business of extending invitations to government, but the consequences of ten or more years of financial restraint in our universities are clear: most have arrived at a strategy for the pursuit of research and teaching where resources are limited. It is a question of emphasis and priority; which will be sacrificed first, research or teaching? That the question is asked should be a matter of concern to everyone interested in higher education in the United Kingdom.

Piecemeal cessation of research activity will only increase disparities among institutions and lead to further decline in UK science. The government should encourage universities with low research grades to do better through increased funding for tertiary level education and public sector research. This is the only realistic means whereby the United Kingdom can expand its graduate output, avoid increasing shortages of scientists and create an economic and cultural environment worthy of a developed country in the twenty-first century. For the time being, many of the poor relations among the UK university research community can take pride in the high demand for their graduates.

W. I. MONTGOMERY

School of Biology and Biochemistry,
The Queen's University of Belfast,
Belfast BT9 5AG, Northern Ireland

More on Shockley

SIR—Frederick Seitz's "Defence of Shockley" (*Nature* **342**, 474; 1989) omits events equally as important as the traumatic automobile accident on 23 July 1961 "in evaluating his activity in later years" — "his . . . ill-conceived concentration on socio-genetic matters": His marriage, which had produced three children, was dissolved. John Bardeen told me (December 1966) that Shockley "tried to solve all the problems in physics, had a nervous breakdown in 1951" and entered a hospital for psychiatric care¹. As reported in a 1962 biography of Shockley (based on interviews with him) ". . . there was a year or two when Shockley's greatest scientific interest was psychiatry". When he subsequently remarried, it was to a woman who "has taught psychiatric nursing both in hospitals and at Ohio State University"².

JERRY HIRSCH

University of Illinois at
Urbana-Champaign,
603 E. Daniel Street,
Champaign, Illinois 61820, USA

1. Hirsch, J. *SAGE Race Relations Abstracts* **6**, 1–65 (1981).
2. Thomas, S. *Men of Space: Profiles of the Leaders in Space Research, Development, and Exploration* **4**, 195 (Chilton, Philadelphia and New York, 1962).

Open doors for research in Japan

Thomas W. Ebbesen

Many Japanese companies are keen to recruit foreign scientists, and as employers they have much to offer. More researchers should take up the opportunities available for work in Japan.

THE difficult research conditions in many of Japan's universities have been discussed in recent Commentaries^{1,2}; no doubt such circumstances discourage foreign scientists from seeking employment in Japan's academic institutions. But more than 80 per cent of Japan's research and development is carried out in the private sector where, because of the phenomenal growth of Japanese industry, the R&D environment is evolving rapidly in terms of funding, direction and hiring practices. There are challenging opportunities here for foreign researchers.

Just as public institutions in Japan now have many postdoctoral fellowships available, many companies have also opened their doors to foreign scientists. A survey conducted by the Tokyo office of the US National Science Foundation in 1985 showed that more than 60 per cent of the responding companies were either ready, or would be ready, to employ foreign researchers in their laboratories³. Although it remains to be seen whether these plans have been translated into action, it is clear that many companies now have foreigners on their staff. Companies believe that recruitment of foreigners supports the trend towards increased basic research by providing fresh input from people with different perspectives and backgrounds.

In general, the number of foreigners working in the private sector, both academic and industrial, is much greater than in the public sector. In 1987, for instance, there were 392 foreign full professors in private universities and 12 in national and public institutions⁴. The reason for this large difference is probably that public institutions have not provided a general mechanism for tenure of their foreign employees⁵. As a result, most of them wishing to stay in Japan end up with a job in a private university or in industry. Among the more progressive companies, both temporary and permanent employment are provided. In 1988, for instance, among the 522 foreign graduates of Japanese universities who were hired by Japanese companies, more than 40 per cent had permanent posts⁵. This large percentage is very encouraging as the offer of permanent employment to these foreigners who have little trouble with Japanese suggests openness and fair treatment irrespective of nationality.

Like the public sector², the private sector rarely advertises jobs in journals. Traditionally, recruiting is done through direct contact, a human network, between

university professors, company employees and the prospective candidates. In general, recruiting is not as competitive as it is, for instance, in the United States. Japanese companies rely heavily on their reputation and desirability to attract the best graduates. Foreign scientists are also recruited through direct contacts or by recommendation by a third party. Human relations being of the utmost importance to the management, evaluation of the person both as a 'social being' and as a scientist has a high priority.

Foreign researchers are generally hired on yearly renewable contracts such as postdoctoral positions. Even in the companies that offer permanent employment, there is generally a period of adaptation and mutual evaluation before an employee achieves permanent status. Permanent employment requires strong long-term commitment by both parties.

There is a general belief that Japanese institutions have a much more rigid hierarchy than those in other countries. But this depends on the type of institution. Company laboratories can be quite democratic: the difference in salary between the lowest and highest ranks is small, and regular meetings are held at all levels to exchange information and discuss policies. Managements have their systems of checks and balances. For instance, a lower-ranking researcher can request a transfer to another group if he or she is dissatisfied with working conditions. So it is important to managers to have a good relationship with their subordinates and colleagues, and decisions are generally taken by consensus and compromise. Although some complain that such a system does not allow for clear-cut and quick decisions, it does mean that human relations are smooth.

Because of this system, the period between a potential employee's first contact with the company and his or her hiring can be very long. It takes at least 5 to 6 months to process an application and for visa formalities. Furthermore, foreign candidates unfamiliar with Japanese practices often find it difficult to judge the true level of commitment from their correspondence with a company. Although knowledge of Japanese is generally not essential for postdoctoral or visiting status, some knowledge of the Japanese language and culture is very useful and can decrease cross-cultural barriers which are mutually frustrating.

As was pointed out by Yamamoto¹, the

overall expenditure on research by industry is much larger than that by the public sector (although spending on basic as opposed to applied research remains to be considered). The difference is even greater when high-tech industry is considered: for example the electronics industry obviously depends on innovation and improvement, which in turn requires significant spending on research and development. The average amount per researcher spent on capital equipment in the larger company laboratories can only be matched by the most powerful research groups (*koza*) of public institutions, and is comparable to that available in the equivalent top Western companies.

With the favourable financial position of many companies in Japan at the moment, research investment is becoming more stable, long-term and more orientated to basic science. In a few companies, research is being done that is decades ahead of development. Companies offering such opportunities are attractive alternatives to an academic career, generally considered more prestigious by young Japanese researchers. These conditions are also attracting an increasing number of foreigners.

There is a problem which hinders scientific exchange. Young postdoctoral scientists, especially from the United States, are often neither encouraged to go abroad nor credited for their scientific experience there, so postdoctoral positions in Japan allocated to US researchers are often undersubscribed despite the efforts of agencies like the US National Science Foundation and the Science and Technology Agency of Japan². It is hard to understand the reasons for this attitude as foreign companies are building R&D facilities in Japan to pool the local resources, and hire Japanese scientists from rival companies. Complaints about the closed nature of Japan's research are no longer appropriate; rather it is time to encourage researchers to take advantage of the many opportunities now on offer in Japan's private and public sectors. □

Thomas W. Ebbesen is in the Fundamental Research Laboratories of NEC Corporation, Tsukuba 305, Japan.

1. Yamamoto, A. *Nature* **339**, 575-576 (1989).

2. Siddiqui, S.S. *Nature* **340**, 337-338 (1989).

3. *National Science Foundation Tokyo Office Report Memorandum No. 86* (1985).

4. *Gakko Kihon Chosa Hokokusho* (Ministry of Education, 1987).

5. Ministry of Justice Report on Foreign Student Employment (in Japanese) (1989).

Saying sorry, but not sorry

The most demanding art form in the business is that of replying in print to criticism of work already published. By demonstration, practitioners of all disciplines seem to have arrived at similar techniques.

PUBLICATION may not be the unalloyed benefit it is supposed to be by those observers of the research profession who repeatedly remark that a dozen publications in refereed journals count for at least one rung on almost any promotion ladder. Publication also gives hostages to fortune. Others may, for example, write in with criticisms, which is at best a nuisance and may even be a great embarrassment. Is there a recipe, other than taking great care in advance, for avoiding trouble of this kind?

One defence is to publish in journals that decline to publish comments on what may already have appeared, but this is likely to be a shrinking refuge as time goes by. The Popperian doctrine that falsification may be virtuous seems to have made a mark, even if a little off target, and the intelligence that even careful authors make mistakes appears to be spreading. Journals considering themselves immune from error and thus above the fray are a declining company. So will there be no escape from criticism?

Nobody should be dismayed. The journals that already publish comments and rejoinders (this one included) are replete with illustrations of how it is possible to reply with dignity, and all the appearance of having been in the right all along, to the most hostile and even cogent criticism. The examples in what follows have been anodysed and bowdlerized where necessary to conceal the identity of the authors and their journals. They are taken from journal issues some years back, so that passions will have cooled even when the authors recognize their own prose. That seems only fair when the practice of saying sorry without saying so is so nearly universal.

Those replying to criticisms must first choose one of the several different styles in which a rejoinder may be couched. One much favoured style is the haughty, in which the responder puts his critics firmly in their place with a patronizing generalization of the complaint.

Thus there is a complaint that runs "...this result is a trivial consequence of differential geometry...the experiments reported merely confirm...suggestions that the effect is best understood quantum mechanically are thus misleading ... recent interest raised by the experiment is unwarranted." How can anybody deal with that? Quite simply, it appears. Like this: "One often encounters an ambiguity

in the interpretation of a phenomenon: is it classical, or is it quantal? ... In our experiment, this ambiguity of interpretation also arises..." There follows a series of three ripostes, all slightly oblique to the complainant's arguments.

The haughty rejoinder style may naturally be elaborated as the responder's imagination will allow. It makes natural sense to remind a critic that the original objectives of the published paper were much broader, even grander, than the puny objections now raised, as with a phrase such as "We were concerned with the implications of a direct connection between and energetically dominated thermodynamic ensembles. This theory provides a variety of specific predictions... XY&Z say that their experiment directly contradicts one of the predictions, but that conclusion depends on unsubstantiated assumptions."

The other most common defence against criticism is the retreat into complexity, which again is most often introduced with a few haughty sentences. Thus the last rejoinder continues with the assertion that the critics have failed to demonstrate that assumptions about the temperature of two phases are correct and that they have not allowed for changes of the mobility of certain carriers in a certain class of solids. It seems common, and no doubt effective, practice to cast aspersions on the critics' competence with references to reliance "on experiments of others" and host of similar remarks.

Haughtiness can also be combined with a defence that is the opposite of complexity — the position that the data available do not justify the intricacy of the critics' arguments, as in the sentence "We believe it would be a mistake to over-interpret the existing data." If he is but partly human, it is the critic and not the responder who will have the impression that he has made a fool of himself. The general principle (for responders) is that one should never openly say one has been mistaken.

But what if the circumstances are such as to require making a clean breast of error? Even then one can retain a sense of dignity by an appropriate choice of words. Thus, "We did not intend to rely too naively on the ... idea, though we admit that our ideas are evolving on this matter" or "We agree with these two points and will try to clarify the misunderstanding that led to the criticism, point out some problems in it and correct a concep-

tual error in the analysis of the data ... which, however, has no influence on its conclusion".

The conclusion, apparently designed to exonerate everybody concerned from all imaginable errors, is that "One conclusion that should be drawn from this controversy is that theoretical models should not be tied too closely to the experimental systems, which are usually much more complex. Another conclusion is that a theoretical model may stimulate the understanding of experimental observations even though the specific assumptions on which the model is based are not fulfilled in the experiment."

Again, it seems, there is no limit to authors' ways of acknowledging they have been given pause without acknowledging anything so mundane as an error. Thus "X *et al.* raise an important point They then proceed to show that in the experiment of Y *et al.* the parameter is too large by a factor of three... We agree with their results in essence (but see below)..." This, as it turns out, is a trick for blaming the whole misunderstanding on hapless Y *et al.*, the sources of the disputed data.

Further recitation of these forms of words could only engender cynicism about the reality of the open debate about the rights and wrongs of important issues, when there are important procedural issues underlying these exchanges. The plain truth is that people in the same field as the original authors and their critics are usually quick to tell who is right and who is wrong, but that people in other fields must go to a great deal of trouble before they can find out. The result is that nobody's pride is publicly injured, and that the record is set straight among those in the know. The losers are those who read the journals for general instruction.

The remedy, curiously enough, exists already. Complaints about a published article are usually referred not merely to the original author, but to those who reviewed that publication, in the hope that they may spot a pointless misunderstanding, and who may themselves consider they have some explaining to do. There is a sense in which the eventual exchange evolves from the frequently indignant content of the complainant's first covering letter to the minuet-like formality of what eventually sees the light of day. Perhaps more unvarnished protests, and their authors' reputations, should be put directly into hazard.

John Maddox

Precursors to eruption

Wayne Thatcher

JAPANESE scientists have assembled a uniquely detailed picture of events that led up to the submarine eruption of Teishi volcano, off the Izu Peninsula 100 km southwest of Tokyo¹. The eruption was small and was not destructive but data gathered by the Japanese illustrate the merits of joint monitoring of seismic activity and crustal deformation for assessing volcano hazard. On page 631 of this issue², Shimada *et al.* describe the role played by the new satellite-based Global Positioning System (GPS) in monitoring Teishi and point out the great potential of this method for measuring crustal deformation and learning about how volcanoes work.

Volcanoes are built by the repeated eruption of hot fluid rock (magma) and rock fragments onto the Earth's surface. Magma and associated gases reside beneath the volcanic edifice in subsurface reservoirs (see figure). Physical and chemical changes in the cooling and solidifying magma cause pressure changes in and around the reservoir that lead to lateral and upward migration of magmatic fluids through new or reactivated conduits. This migration can often be detected by locating small earthquakes triggered by the stress changes accompanying magmatic injection. Pressure changes in the magma reservoir and stresses induced by magma and fluid migration also result in deformation of the Earth's surface that can be monitored by geodetic methods and by continuous measurement of ground tilt or strain at the surface or in shallow boreholes.

Volcanoes are only intermittently active and are typically dormant for long periods between eruptions. Although small earthquakes and crustal movements are clear indicators of magmatic unrest, eruption is by no means an inevitable consequence. Indeed, the great majority of these episodes pass with neither a bang nor a whimper³. During the past decade, magmatic resurgence at Rabaul (Papua New Guinea)⁴, Pozzouli (Italy)⁵, Long Valley (California)⁶ and the Izu region have made these among the world's most closely watched volcanoes, but of these only the Izu volcanoes have erupted thus far.

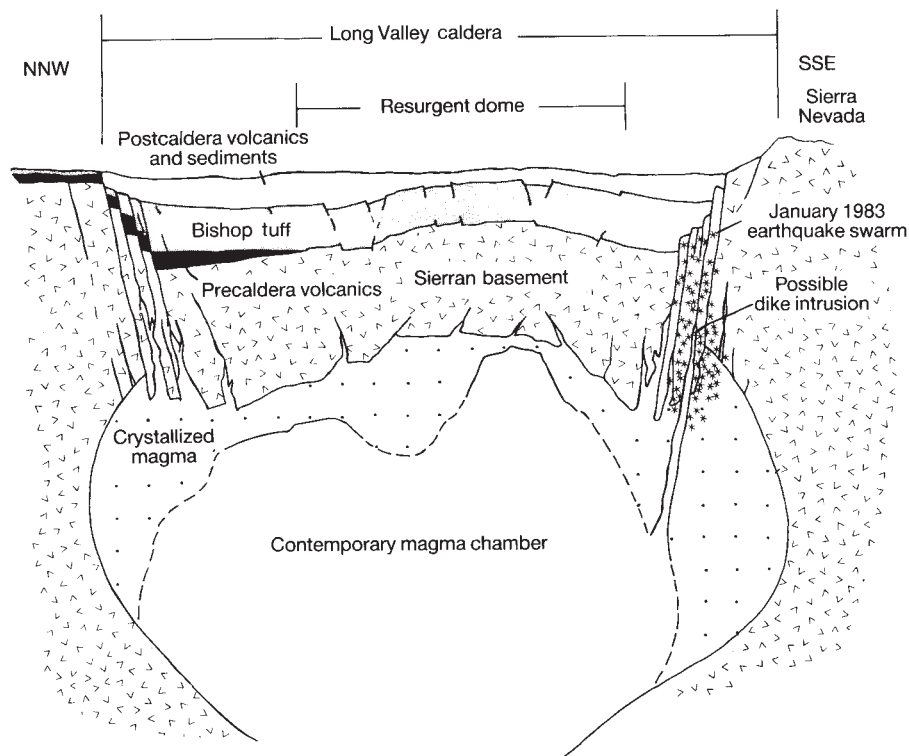
The Izu Peninsula experienced intense swarms of tremors and domal uplift during 1925–32 and a magnitude-seven earthquake in 1930, but no eruptions occurred. After more than 40 years of dormancy, crustal movements and earthquake swarms resumed in 1975 and continue to the present. The recent uplift occupies a broad onshore region about 40 km in diameter which undoubtedly

also extends significantly offshore. Magnitude-seven earthquakes occurred on the periphery of the uplifted region in 1978 and 1980 and simple model calculations suggest that they were triggered by the pressurization of a magma reservoir centred about 10 km beneath the maximum uplift⁷.

The sequence of events that led up to the Teishi eruption is one of the most complete case histories ever documented. The centre of uplift on the Izu Peninsula has shifted slightly eastwards since the late 1970s, roughly towards the site of Teishi volcano. However, a critical premonitor of eruption is not the subsurface movement of magma but its ascent to the Earth's surface. Again, seismicity and crustal deformation supplied the information needed to locate the zone of upward migration and to follow its temporal evolution.

The region around Teishi volcano has shown intermittent seismic activity since late 1978, with 12 distinct earthquake swarms including a sequence in 1980 with a magnitude-seven event. An intense swarm that started on 3 July 1989 differed from those preceding it in that it was accompanied by large and rapid crustal movements recorded by many instruments, including GPS receivers. The movements and the swarm activity, interpreted by Shimada *et al.*² as an upward propagation of a magma conduit, continued until 10 July. Ground tremors and strong rumbling sounds observed on 11 and 12 July may have been due to submarine eruptions that were not witnessed; on 13 July, updoming of the sea surface and a large ash cloud provided clear evidence of eruption.

Crustal movements involved in magma migration in the upper crust are typically quite large, involving strains and tilts of 10 parts per million (p.p.m.) or greater, well within the resolution of both conventional geodetic survey techniques and continuous strain monitoring. Common-



Long Valley caldera in eastern California has experienced magmatic unrest since 1978 and, along with the Izu Peninsula, is one of the world's most intensively monitored active volcanic regions. The cross-sectional view of the caldera shown here (from ref. 10) represents a synthesis of geological, seismic and geodetic information into a picture of subsurface structure that is representative of many active magmatic systems. The magma reservoir is viewed as a volume containing a heterogeneous distribution of molten rock lying below 5 km depth. Since the late 1970s, pressure changes within this region initiated the current period of unrest and resulted in earthquakes as large as magnitude six and ground uplift of 0.5 m. An intense seismic swarm (asterisks) and large amounts of crustal extension near the southern rim of the caldera measured in January 1983 are believed to have been a result of upward migration of magma to within about 3 km of the surface. This episode was similar to the July 1989 Teishi activity, but it did not culminate in an eruption. Since 1984, crustal deformation rates and seismicity have declined but the region remains restless. Since May 1989, a localized swarm of earthquakes has persisted beneath Mammoth Mountain, a ski resort near the southwest rim of the caldera, and this region is still being closely watched.

ly, however, logistical constraints and cost limit volcano surveillance and here the new technology of GPS has a potentially important contribution to make.

GPS surveying relies on interferometric correlation of radio noise transmitted by Earth-orbiting satellites and simultaneously recorded by receivers located at geodetic benchmarks on the Earth's surface⁸. Comparison with ground-based measurements shows that GPS methods can determine relative horizontal positions within 0.4–0.7 p.p.m. over baselines 10–40 km in length⁹. Although these capabilities are at present a factor of two to three less precise than the best conventional methods, GPS has a number of advantages over ground-based surveying. Because the GPS receivers track radio signals from orbiting satellites, ground stations need not be visible to one another and clear skies are not required. Tracking can be pre-programmed into the receivers and observing is straightforward. In contrast, high-precision ground surveying is an exacting, time-consuming skill that is manpower intensive and requires long-term commitment. Until recently most of this surveying had been done only by large organizations such as national geodetic surveys. GPS is changing all this, and the field of crustal deformation measurement is now accessible to a wide spectrum of researchers whose results will have great importance for seismology, tectonics and volcanology.

New observation networks have only recently been established and the full impact of GPS has yet to be felt. Most of the new measurements are being made at active plate boundaries, where deformation rates average only about 0.3 p.p.m. per year and 5–10 years in intersurvey movement is needed to determine a well-constrained deformation field. More immediate results can be obtained by resurveying existing geodetic control networks, many of which were established at the end of the nineteenth century, but this source of data has been less fully exploited.

Paradoxically, one of the most promising applications of GPS, volcano monitoring, has scarcely been explored. As Shimada *et al.* show, the detection of deformation caused by magmatic processes is well within the measurement

capabilities of GPS. Furthermore, the signals observed near Teishi volcano were so large that the time-consuming post-processing of data required accurately to determine satellite orbits and to extract the maximum precision attainable by the method was unnecessary. Consequently, the Japanese investigators could take advantage of the capability of GPS for continuous unattended recording and obtain daily measurements of changes in baseline length.

The eastern Izu Peninsula is one of the most intensively monitored regions on

COSMOLOGY

Kinky cosmic strings

Alexander Vilenkin

CERTAIN theories of elementary particles predict the existence of thin lines of concentrated energy called cosmic strings. Such strings could be formed at a phase transition shortly after the Big Bang and may still be present today in distant regions of the Universe. The great interest enjoyed by strings among cosmologists is due, among other things, to their possible role in the formation of galaxies and cosmic large-scale structure. Evidence, such as gravitational radiation, to confirm the existence of cosmic strings has been hard to find. New calculations^{1,2} of the way that they might have evolved since they formed could provide an explanation for this absence of evidence. They also must modify our ideas of how strings could affect the structure of the Universe.

The mass per unit length required for strings to cause galaxy formation is about 10^{22} g cm⁻¹. This is in the range expected from grand unified theories of elementary particles. More significantly, the gravitational field of strings with this density would be sufficiently strong to produce a number of potentially detectable observational effects: double images of distant galaxies and quasars as a result of gravitational lensing; linear discontinuities in the intensity pattern of the microwave background radiation; and a broad spectrum of low-frequency gravitational waves.

When strings are produced at a phase transition, they form a tangled network permeating the entire Universe. They do not have ends and can either form closed loops or extend to infinity. The evolution of this tangle of strings is rich in physical processes. Tension in curved strings makes them wiggle violently and they often cross themselves and one another. Intersecting strings break at the crossing point and reconnect the other way. Long, twisted strings cross themselves many times and shed their coils in the form of closed loops. The loops can self-intersect

Earth and GPS was one of many methods used to measure crustal movements, so its capabilities were by themselves not crucial to the observing programme. In many volcanically active regions, however, monitoring is limited or nonexistent and conventional surveying is not feasible. GPS methods combined with simple seismic observations have great potential for supplying the kinds of data needed in successful eruption forecasting. □

Wayne Thatcher is with the Branch of Tectonophysics, US Geological Survey, Menlo Park, California 94025, USA.

and further fragment, but eventually the fragmentation ceases and each loop leaves behind a family of non-intersecting daughter loops. As these remaining loops oscillate under the action of their tension, they gradually lose their energy to gravitational radiation, shrink and disappear.

The cosmological evolution of the string network is believed to be scale-invariant: that is, statistically the network looks the same at all times. The only thing that changes is its overall scale, which is set by the horizon length, h , equal to the distance that light has travelled since the Big Bang. The scale-invariant character of string evolution suggests that at any cosmic time t , the distance between nearby long strings is comparable to the horizon, $h \sim ct$ (c is the speed of light), and that the typical size of the loops produced by the network is also a fixed fraction of the horizon.

The most surprising result of the new simulations^{1,2} is that long strings have a significant fine structure on scales much smaller than the horizon. Most of this structure is in the form of sharp angles, 'kinks' formed at the points of string reconstructions, so that the strings have zigzag shapes. Moreover, the typical size l of stable loops is comparable to the scale of the smallest zigzags, $l \ll h$. The kinks on long strings are gradually smoothed out by the expansion of the Universe and by gravitational radiation, and the scale l should be determined by an interplay between the kink production and smoothing. At present we do not have a reliable estimate for l , but the simulations indicate that it does not exceed $10^{-3}h$. Earlier simulations by Albrecht and Turok³ suggested a scaling behaviour of long strings, but did not have enough resolution to detect the fine structure.

The new picture of string evolution has important consequences for the string theory of galaxy formation. There are basically two ways in which the strings can produce cosmic structure. The first is

1. *Sci. Event Alert Network Bull.* **14**, 2–5 (1989).
2. Shimada, S. *et al. Nature* **343**, 631–633 (1990).
3. Newhall, C.G. & Dzurisin, D. *US Geol. Surv. Bull.* **1855** (1988).
4. McKee, C.O. *et al. Bull. Volcanol.* **47**, 397–411 (1984).
5. Barberi, F. *et al. Bull. Volcanol.* **47**, 175–185 (1984).
6. Rundle, J.B. & Hill, D.P. *A. Rev. Earth planet. Sci.* **16**, 251–271 (1988).
7. Thatcher, W. & Savage, J.C. *Geology* **10**, 637–640 (1982).
8. Wells, D. *et al. Guide to GPS Positioning* (Canadian Geodetic Associates, Fredrickton, 1987).
9. Davis, J.L. *et al. J. geophys. Res.* **94**, 13635–13650 (1989).
10. Hill, D.P. *et al. J. geophys. Res.* **90**, 11111–11120 (1985).

through the gravitational attraction of matter to oscillating loops; the second is through the formation of wakes behind rapidly moving long strings. In the old picture, which assumed that the loop sizes were set by the horizon, these two effects were expected to be of comparable magnitude. Now, we know that the loop sizes are much smaller than the horizon, which implies that the role of wakes is much greater than that of the loops. Allen and Shellard² made an interesting observation that the coherent velocities of long strings in the network are rather small ($v \approx 0.15c$), with the exception of the highly curved string segments affected by recent reconnections which tend to have much higher velocities. The wakes produced by such rapidly moving segments are particularly strong and may lead to the formation of very large mass concentrations, such as the Great Attractor. A quantitative comparison of the string theory with observations should await the results of numerical simulations, currently underway, that combine the string-evolution computer code with a simulation of the gravitational clustering of matter.

Finally, what do the new simulations imply for the gravitational-wave background expected from cosmic strings? The observational bounds on the intensity of low-frequency gravitational radiation have been recently improved. The best bound comes from the millisecond-pulsar timing measurements by Taylor⁴: gravitational waves should disturb the regular pulse trains. Taylor's preliminary results⁴ give an upper limit that is too small for the theory based on the old picture of string evolution⁵⁻⁸. But the small loop sizes obtained in the new simulations considerably reduce the predicted intensity of radiation. With new results taken into account, Taylor's bound can be expressed⁹ as an upper bound on the string's mass per unit length, $5 \times 10^{22} \text{ g cm}^{-1}$. This falls short of imposing a significant constraint on the string model by an order of magnitude. But with both theory and experiment developing so rapidly, a direct confrontation between them will be inevitable in the near future. □

Alexander Vilenkin is in the Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA.

Electrochemical promotion

J. Pritchard

THE extent to which the catalysis of reactions at a surface is essentially a short-range effect, depending on the character of specific surface sites, or also involves long-range electronic interactions, is controversial. The report by Vayenas *et al.*, on page 625 of this issue¹, appears to show that long-range influences can indeed be important.

Advances in surface science during the past two decades have transformed our understanding of the nature of solid surfaces and the way that molecules from the surrounding medium interact with them. Nevertheless, heterogeneous catalysis continues to pose many problems. Normal operating conditions for catalysts are far removed from the low pressures required for most surface-science techniques, and *in situ* techniques are being developed to enable the direct study of model surfaces to be extended from the usual ultrahigh vacuum regime to realistic working conditions. Most process catalysts are also complex materials that have been empirically optimized in composition and preparative procedures. Supported metal catalysts, in particular, usually contain additives, often alkali metal compounds, that enhance activity or improve the selective yield of a particular product. Understanding the mechanisms of such promoters through studies of model systems is an aim of much current research by the methods of surface science².

The electrochemical modification of metal catalyst surfaces described by Vayenas *et al.* has the particularly attractive features of permitting continuous *in situ* variations and monitoring of the amount of the promoter. The approach the authors developed uses porous metal-film catalysts deposited on the surface of a solid electrolyte which enables oxygen ions or sodium ions to be transferred to the catalyst surface. Such ions modify the work function (ionization energy) of the metal and change its catalytic activity, but their role is not that of an electrochemically provided reactant: the increases in the rate of ethene oxidation over platinum, for example, are orders of magnitude greater than the rate of electrochemical supply of oxygen ions, and the surface is already largely covered by oxygen adsorbed from the gas phase. The rate changes are accompanied by changes of the apparent activation energy which show, over a wide range, a one-to-one correlation with changes in the work function, also measured *in situ* with a Kelvin probe. Conversely, sodium ions produce a decrease in work function and a reduction in the rate of reaction.

A long-standing question is whether

promotion and poisoning by co-adsorbed electronegative or electropositive species are essentially short-range effects involving sites adjacent to the co-adsorbate or long-range effects mediated through the catalyst's electronic band structure. Recent theoretical studies³, supported by experiment⁴, show that short-range effects dominate, strongly affecting bonding to surface atoms only within a range of less than 0.5 nm, in contrast with the new results of Vayenas *et al.*

For example, the importance of alkali promotion in the synthesis of hydrocarbon, alcohols and so on, by hydrogenation of carbon monoxide has led to very detailed investigation of the nature of the interaction of carbon monoxide with alkalis on single-crystal surfaces of several transition metals. The local interaction can be very strong, leading in some cases to coincident thermal desorption of CO and potassium, for example. The formation on copper surfaces of a distinct compound of CO with potassium has been proposed⁵, though in the strong interaction with potassium on ruthenium and palladium surfaces^{6,7}, the CO is believed to remain bound to the transition metal. At low coverages of potassium, where it is thought to be ionic, the interaction with CO seems to be of longer range. Instead of inducing distinct states that desorb at high temperatures, the effect is to modify only slightly the CO bond to the surface⁷. The CO desorption energy increases by only 15 kJ mol⁻¹, whereas the potassium reduces the work function by about 350 kJ mol⁻¹.

Work-function changes measured by contact potential methods reflect the long-range average electrostatic potential differences between the interior of the metal and points that are far outside the metal compared with chemical bond lengths or the spacing of adsorbed species even in dilute overlayers. At low coverages, the local surface barriers depend very strongly on the lateral distance from the adatom. This has been shown very elegantly by photoemission spectra of adsorbed xenon in which the binding energy of the outer-shell electrons referred to the Fermi level directly reflects the local work function⁸. For potassium on ruthenium, xenon atoms on sites next to

1. Bennett, D. P. & Bouchet, F. R. *Phys. Rev. Lett.* **63**, 2776–2779 (1989).
2. Allen, B. & Shellard, E. P. S. *Phys. Rev. Lett.* **64**, 119–122 (1990).
3. Albrecht, A. & Turok, N. *Phys. Rev. Lett.* **54**, 1868–1871 (1985).
4. Taylor, J. H. *Pap. presented at US–USSR Workshop on High-Energy Astrophys.* (Tbilisi & Moscow, 1989).
5. Vachaspati, T. & Vilenkin, A. *Phys. Rev. D* **31**, 3052–3058 (1985).
6. Hogan, C. J. & Rees, M. J. *Nature* **311**, 109–114 (1984).
7. Albrecht, A. & Turok, N. *Phys. Rev. D* **40**, 973–1001 (1989).
8. Accetta, F. & Krauss, L. *Phys. Lett.* **233B**, 93–98 (1989).
9. Bouchet, F. R. & Bennett, D. P. *Phys. Rev. D* **41**, 720–723 (1990).

1. Vayenas, C. G., Babelis, S. & Ladas, S., *Nature* **343**, 625–627 (1990).
2. Thomson, S. J. *J. chem. Soc., Faraday Trans. I* **83**, 1893–1914 (1987).
3. Joyner, R. W. & Pendry, J. B. *Catal. Lett.* **1**, 1–5 (1988).
4. Caddock, P. D., Joyner, R. W. & Williams, B. P. *Catal. Lett.* **2**, 27–34 (1989).
5. Lackey, D. & King, D. A. *J. chem. Soc., Faraday Trans. I* **83**, 2001–2013 (1987).
6. Kiss, J., Klivenyi, G., Revesz, K. & Solymosi, F. *Surf. Sci.* **223**, 551–568 (1989).
7. Berko, A. & Solymosi, F. *J. chem. Phys.* **90**, 2492 (1989).
8. Markert, K. & Wandelt, K., *Surf. Sci.* **159**, 24–34 (1985).

or above potassium show large shifts, whereas more distant xenon atoms show values much closer to the clean ruthenium value.

The strong long-range effect implied by the correlation of work-function change with activation-energy change found by Vayenas *et al.* in the electrochemically induced promotion² is therefore particularly intriguing. So too is the nature of the electrochemically induced oxygen species that is believed to cause the increase in work function and catalytic promotion, yet which is less reactive than the adsorbed oxygen reactant that covers most of the

surface. There is clearly much surface chemistry to be explored by this technique and it will be interesting to see how general the work-function effect proves to be. In any case, the ability to vary the concentration of promoters by electrochemical control while under reaction conditions is a valuable development in catalytic research, and one can expect it to be rapidly exploited in conjunction with other *in situ* techniques of surface analysis. □

J. Pritchard is in the Department of Chemistry, Queen Mary and Westfield College, University of London, London E1 4NS, UK.

IMMUNOLOGY

Questions of presentation

Peter Cresswell

THE mechanisms involved in antigen processing and presentation are proving to be highly sophisticated. At a recent meeting* it became clear not only how far we have come in understanding the process, but how far we have to go.

The principal function of the major histocompatibility complex (MHC) glycoproteins is to bind peptides from foreign antigens and display them on the surface of cells. The MHC-peptide complexes are then specifically recognized by T lymphocytes. Class I MHC glycoproteins interact with specific receptors on CD8-positive cytotoxic T-lymphocytes, class II ones with receptors on CD4-positive helper T-lymphocytes. Recognition of a specific peptide-class I complex usually results in the death of the cell displaying it. Recognition of a peptide-class II complex induces proliferation of the CD4-positive T cell and amplification of the immune response to the particular immunogen from which the peptide was derived. Normally, class I MHC-restricted T cells kill cells infected by viruses. Because virtually all cell types express class I MHC molecules, any infected cell can become a target for cytotoxic T cells. Class II MHC-restricted helper T cells, on the other hand, respond to B cells, macrophages or other antigen-presenting cells which have internalized exogenous protein antigens. Class I MHC molecules are believed to present peptides derived from cytoplasmic proteins synthesized by the cell, whereas class II MHC molecules generally present peptides derived from exogenous proteins.

Although the protein antigens presented to T cells may be synthesized in the cytoplasm, it is not clear whether the peptides derived from them are generated there or after transport of the intact

protein into the secretory pathway. Townsend and co-workers have shown that modification of the N-terminal amino acid of influenza virus nucleoprotein to enhance its rate of cytoplasmic degradation increases the ability of cells expressing it to be recognized by class I-restricted cytotoxic T cells¹. This supports the hypothesis that the peptides presented by class I MHC molecules are generated by cytoplasmic proteolysis.

Early binding

Once generated, the peptides are thought to bind to class I MHC molecules early in transport. Brefeldin A, which inhibits transport from the endoplasmic reticulum (ER) to the Golgi apparatus, inhibits class I-restricted T-cell recognition of virally infected cells^{2,3}. It has no effect on the lysis of uninfected cells if the appropriate epitope, as a synthetic peptide, is added to cells, when association of peptide with mature class I molecules is presumed to occur at the cell surface. The interpretation of the brefeldin A effect is that class I-peptide complexes are prevented from leaving the ER and being expressed on the cell surface. It was reassuring to hear at the meeting that a more specific means of preventing the egress of class I MHC molecules from the ER also prevents cytotoxic T-cell recognition. Expression of the adenovirus E3/19K protein, which is resident in the ER and binds class I MHC molecules, inhibits H-2D^k- or H-2L^d-restricted killing of influenza- or vaccinia-infected mouse cells (J. Yewdell, National Institutes of Health). These results strongly support the argument that the ER, or an associated early compartment, is where class I MHC molecules bind peptides.

The molecular events involved in peptide binding by class I MHC molecules are far from clear. A recent step forward was the identification of a mouse cell line, RMA-S, which is defective in its ability to

process viral antigens⁴. This cell line expresses low levels of the endogenous H-2^b class I molecules; and, when infected by influenza virus, it cannot be killed by influenza-specific H-2D^b-restricted cytotoxic T cells. It can, however, be killed when the specific epitope is added to the cells as a synthetic peptide. A human B-lymphoblastoid cell line (LCL) .174, which shows a similar defect in class I MHC antigen transport, is also unable to process and present viral antigens to class I-restricted cytotoxic T cells (A. R. M. Townsend, University of Oxford).

Enhanced expression

Both the mouse and human cell lines show enhanced expression of the appropriate class I-restricting element on the cell surface when incubated with a peptide epitope. For example, incubating the .174 cell line with a synthetic peptide derived from influenza virus matrix protein induces enhanced surface expression of HLA-A2, the restricting element expected to bind it. The enhanced expression is a reflection of an early event in transport, most readily interpreted as binding of exogenously added peptide to the class I MHC molecules in a pre-Golgi compartment which results in a conformation appropriate for transport. In the human mutant cell line the processing deficiency is probably due to mutation of an MHC-linked gene. The transport defect in .174 can be complemented by fusion with normal B- or T-LCL^{5,6}. For such a hybrid to express class I HLA antigens, and to be killed by virus-specific class I MHC-restricted T cells, requires the retention of intact, MHC-bearing copies of chromosome 6 from the fusion partner. Because .174 has a homozygous deletion of the class II region of the HLA complex, a gene essential for processing of viral antigens is probably encoded in this region. A key goal is to identify this gene and its mechanism of action.

Internalization and proteolysis of antigens are prerequisites for their class II MHC-restricted presentation to T cells. By chemically coupling cytochrome *c* to monoclonal antibodies that react with different surface molecules of mouse B cells, the processing and presentation of cytochrome *c* can be increased dramatically (S. Pierce, Northwestern University). Such findings are thought to reflect endocytosis of the antibody conjugates by the reactive surface molecules. This may be an oversimplification because internalization of certain molecules which functionally enhance antigen presentation, such as class I MHC molecules, has not been observed in B cells. Surface immunoglobulin molecules, however, are endocytosed rapidly following antigen binding, and specific antigens so

* Cellular Routes of Antigen Processing and Presentation, Cancer Research Institute, New York, 14-15 November, 1989.

internalized are rapidly exposed to proteinases (C. Watts, University of Dundee).

Evidence favouring an endosomal, rather than lysosomal, encounter with proteinases comes from immunoelectron microscopic studies (F. Brodsky, University of California, San Francisco)⁷. Anti-immunoglobulin-antibody-coated colloidal gold particles can be detected in peripheral endosomes within two minutes of binding to the surface immunoglobulin of human B-LCL. Two proteinases, cathepsins B and D, can be co-localized to the endosomes containing the colloidal gold particles. Furthermore, HLA-DR molecules can be detected in the same endosomal compartment. These results support a model for antigen processing, at least in B cells, in which internalized antigens are proteolysed in the endosome, and the peptides derived from them are immediately accessible to class II MHC molecules. These molecules bind a subset of peptides and are then transported to the cell surface.

Recycling?

One of the remaining questions concerns the origin of the class II MHC molecules which bind peptides. Here (as so often in immunology) there are two extreme views — first, that the class II molecules are all derived from the cell surface by endocytosis; second that they are all drawn from a newly synthesized intracellular pool. The first hypothesis proposes that class II molecules are recycled between the plasma membrane and the endosome and that exchange of bound peptides occurs during this process. Some experiments indicate that peptide exchange can occur *in vivo* (L. Adorini, Sandoz Ltd, Basel)⁸. Antigen-presenting cells incubated with a hen egg lysozyme-derived synthetic peptide are recognized by an I-A^k restricted T-cell hybridoma. Co-incubation with a second I-A^k-binding peptide, which does not stimulate the hybridoma, will compete with the antigenic peptide and inhibit its recognition. On fixed cells, or cells cultured at 18 °C, the peptides must be added simultaneously for effective competition. With live cells at 37 °C, adding the competitor peptide even eight hours later results in inhibition, indicating that the competitor peptide is being exchanged for antigenic peptide associated with a pre-formed peptide-I-A^k complex.

Corroborating experiments in another system show that the dissociation of class II-associated peptides is rapid in live B cells, but virtually undetectable in fixed B cells (C. Harding, Washington University, St Louis). Lysosomotropic agents, such as chloroquine, which neutralize acidic, intracellular compartments, have no effect on peptide exchange; so it

seems that pH changes, as such, are not involved. So far there is no direct evidence that recycling of class II molecules is essential for peptide exchange, but it is a reasonable hypothesis upon which to base further work.

If an active process mediates the exchange of peptides, class II molecules may have several opportunities to capture peptides. The second view of the origin of endosomal class II molecules allows no such economy. If the class II molecules entering the endosome are derived exclusively from a newly synthesized intracellular pool, then each would have only a single opportunity to associate with a peptide before its expression on the cell surface.

A characteristic of intracellular class II molecules is their association with an additional glycoprotein, the invariant (I) chain. This glycoprotein associates with class II molecules in the ER and remains associated during transport through the Golgi apparatus. Although low levels of I chain on the cell surface of B-LCL can be detected with certain antibodies (N. Koch, German Cancer Research Center, Heidelberg; V. Quaranta, Scripps Clinic), most of the I chain is proteolytically cleaved and released from class II molecules before expression on the cell surface⁹. In addition to HLA-DR molecules and proteinases, I chain is also detectable in B-LCL endosomes containing internalized anti-immunoglobulin coated colloidal gold particles. The difficulties in demonstrating endocytosis of HLA-DR glycoproteins in B-LCL, together with earlier, indirect evidence that intracellular HLA-DR-I-chain complexes are accessible to the endocytic pathway¹⁰, strongly indicates that most of the endosomal HLA-DR molecules come from a newly synthesized pool rather than the cell surface.

Two principal hypotheses for the function of the I chain have been put forward: that it prevents class II MHC molecules from binding peptides prematurely, perhaps at the point in transport where class I molecules interact with endogenously generated peptides; or that it affects the transport of class II molecules in a manner promoting their endosomal interaction with internalized protein antigens. Evidence for the first hypothesis, from my own laboratory, is that purified HLA-DR-I-chain complexes cannot bind a synthetic influenza virus haemagglutinin epitope, whereas mature, I-chain-free HLA-DR molecules do so efficiently. Further, HLA-DR molecules released from associated I chain by mild denaturation acquire the capacity to bind the peptide. There is no direct support for the second hypothesis but it has been reported¹¹ that I-chain association has no effect on the rate of transport of mouse class II MHC glycoproteins.

Both hypotheses for the function of the I chain predict that class II-positive cells expressing it should process and present antigen to T cells more efficiently than cells that do not. In one report¹² it has been suggested that this is indeed the case. Some experiments indicate that I-A^d-positive cells lacking I chain present peptides better than cells that express it (J. Miller, University of Chicago). An attractive interpretation of these data is that class II molecules transported to the cell surface without the I chain are less likely to be occupied by peptides, and can more easily take up exogenously added ones. The conformation of the surface I-A^d molecules synthesized without the I chain appears to be unusual, because certain anti-I-A^d monoclonal antibodies bind them with considerably diminished affinity.

There is a possible role for a heat-shock protein in class II MHC-mediated antigen processing (S. Pierce, Northwestern University). The linkage of the HSP70 gene to the MHC¹³ and the peptide-binding capacity of heat-shock proteins¹⁴ make them good candidates for study. Mutant human B-LCL, which can present peptides to T cells in a class II-MHC-restricted fashion, but cannot process protein antigens to generate functional peptide—class II complexes, provide evidence for the involvement of additional gene products in antigen processing¹⁵. That cytoplasmic proteins can be presented by the class II-mediated pathway¹⁶ (E. Long, NIH) suggests further complexities. Immunology, which has provided elegant experimental models for geneticists and developmental biologists, is now performing the same service for cell biologists. □

Peter Cresswell is in the Department of Microbiology and Immunology, Duke University School of Medicine, Durham, North Carolina 27710, USA.

1. Townsend, A. et al. *J. exp. Med.* **168**, 1211–1224 (1988).
2. Yewdell, J.W. & Bennink, J.R. *Science* **244**, 1072–1075 (1989).
3. Nuchtern, J.G., Bonifacio, J.S., Biddison, W.E. & Klausner, R.D. *Nature* **339**, 223–226 (1989).
4. Townsend, A.R.M. et al. *Nature* **340**, 443–448 (1989).
5. DeMars, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8183–8187 (1985).
6. Salter, R.D., Howell, D.N. & Cresswell, P. *Immunogenetics* **21**, 235–246 (1985).
7. Guagliardi, L.E. et al. *Nature* **343**, 133–139 (1990).
8. Adorini, L., Appella, E., Doria, G., Cardinaux, F. & Nagy, Z.A. *Nature* **342**, 800–803 (1989).
9. Blum, J.S. & Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3975–3979 (1988).
10. Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8188–8192 (1985).
11. Simons, S., Miller, J. & Cullen, S.E. *J. Immun.* **143**, 3619–3625 (1989).
12. Stockinger, B. et al. *Cell* **56**, 683–689 (1989).
13. Sargent, C.A., Dunham, I., Trowsdale, J. & Campbell, R.D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1968–1972 (1989).
14. Flynn, G.C., Chappell, T.G. & Rothman, J.E. *Science* **245**, 385–390 (1989).
15. Mellins, E. et al. *Nature* **343**, 71–74 (1990).
16. Nuchtern, J.G., Bjeldsson, W.E. & Klausner, R.D. *Nature* **343**, 74–76 (1990).

Case of the diffusing squirrels

John Lawton and Charles Godfray

WITH the benefit of hindsight, the passion of Victorian landowners and naturalists for introducing exotic species into Britain was disastrous. Nowhere is this apparently more clearly illustrated than by the devastation wrought on native red squirrels (*Sciurus vulgaris*) by the American grey squirrel (*S. carolinensis*). But was the grey squirrel solely responsible? Okubo and colleagues¹ now provide a fresh approach to this long-running case.

Grey squirrels were liberated at two main sites around the turn of the century, one at Woburn Abbey in Bedfordshire and another, smaller release, in Yorkshire. Since then, red squirrels have steadily declined in numbers; they no longer occur in most of England, Wales and parts of Scotland² (see graph), and healthy populations are now largely confined to coniferous woodland. But although it seems obvious that the decline of red squirrels is directly attributable to invading greys, the evidence is circumstantial and unclear. A good barrister, defending the accused, could point to two facts. The mechanism by which grey squirrels have exterminated reds is not understood³; there is no smoking gun. Second, at least in some instances, data support the idea that red-squirrel populations collapsed before grey squirrels invaded their former territory^{2,3}. All the innocent grey squirrels have done, your honour, is move into vacant housing.

There is no question that the detailed pattern of replacement of the red by the grey squirrel is complex. Work by Reynolds² in eastern England, for example, apparently confirms that red squirrels became locally extinct before the establishment of their American cousins in

some areas. But in others, the two species co-existed for at least 16 years. Nor did the presence of grey squirrels *per se* enhance the probability of red-squirrel extinction, when examined in over 200 individual 5×5-km squares of the National Grid. However, as Reynolds himself points out, presence or absence of grey squirrels is a very crude measure of effects on the red; a more sophisticated test would seek evidence of density-dependent effects, and such data are not

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The red squirrel — a native under pressure. available. Casual field observations may also be insufficient to detect both species at low densities, giving a misleading impression of what happened in a fine-scale reconstruction of the invasion. Yet despite inadequacies in the data, the consensus among squirrel ecologists today is that grey squirrels do have a severely adverse effect on the abundance of the native red squirrels⁴.

Interestingly, the picture becomes clearer when examined less closely. Although fine scrutiny reveals patchy establishment of greys ahead of the general horde, to a good first approximation the spread of the invader looks like a simple diffusion process, with grey squirrels spreading on an advancing front at an average rate of about 7.7 km a year in East Anglia^{1,5}. Modelling the spread of grey squirrels as a reaction-diffusion process⁶ throws new light on their interaction with red squirrels and increases the likelihood that grey squirrels are guilty.

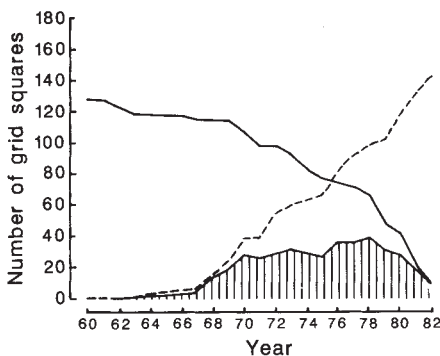
Theories of the spread of an invading organism rest on a seminal paper by Skellam⁶. For an exponentially growing population invading an infinitely large homogeneous environment, Skellam pre-

dicted that the rate of spread of the invading front would approach a constant value. Invading populations of musk rats in Europe^{5,6} and starlings in North America⁷, and recolonizing populations of sea otters in California⁸ have all spread much as Skellam predicted. Now, Okubo *et al.*¹ have developed a model to predict the spread of a population invading a territory occupied by a competitor. The new model differs from Skellam's, and from later extensions and elaborations^{9,10}, by explicitly incorporating interspecific competition as a Lotka-Volterra process leading to a pair of differential reaction-diffusion equations, one representing the invader, the other its resident competitor.

The pair of equations contain eight parameters, four of which can be estimated with reasonable confidence to fit the biology of red and grey squirrels, namely the intrinsic growth rates at low densities and the ultimate carrying capacities of woodlands for each species. The remaining parameters are more difficult to estimate. In order to calculate the rate of advance of the grey and the collapse of the red-squirrel population, we must know the magnitude of the competition coefficients measuring the *per capita* effect of greys on reds and reds on greys. Data to estimate these parameters are not available but it seems reasonable to suppose that intraspecific competition is more important than the effect of red squirrels on grey. If this assumption is accepted, then the rate of advance of the grey squirrel is relatively insensitive to the precise values of the parameters chosen. The authors assume that reds have very little effect on greys (that is, the interaction is essentially amensal, the commonest form of competition¹¹) and that the competition coefficient for greys on reds is 0.15, a very modest impact that would be difficult to detect by field observations.

Estimation of diffusion coefficients is more difficult. Okubo *et al.* explore several approaches to this problem. Estimates based on the observed movements of individual animals indicate a much lower rate of spread than observed. A less direct approach assumes that movement to adjacent, uncolonized woodlands occurs when a habitat becomes saturated by squirrels. The precise value of the diffusion coefficient then depends on the distance between adjacent woodlands. If this distance is assumed to be 10 km, rates of spread consistent with observations are obtained from the model. In general, the estimation of the diffusion coefficient is the hardest part of setting parameters in biological diffusion models.

Simulations of the model generate



Total number of 5×5-km grid squares occupied by red (—) and grey (---) squirrels in East Anglia between 1960 and 1982. The shaded area represents squares shared by both species. (Redrawn from ref. 2).

rates of spread of the grey squirrel consistent with those observed in the field. Moreover, the qualitative results that emerge, such as the pattern in which the grey squirrel spreads from small initial foci that gradually coalesce to wipe out the native red, are highly realistic. Okubo *et al.* conclude that the spread of a species overcoming another in competition is not very different qualitatively from the spread of one in a competitor-free world, except that the wave of advance may be slower. Moreover, simple diffusion, logistic population growth and some general form of rather mild competition are sufficient to account for the progressive replacement of red squirrels by grey squirrels in England and Wales.

Although it may be possible to contrive a model in which red-squirrel populations collapse for reasons that have nothing to do with the presence of alien greys, and in which greys then expand to fill the space available, such a model would

appear to be intrinsically less likely and certainly more complex. Although the jury may still be out on exactly how the grey squirrel did it, it seems increasingly likely that the verdict will be 'guilty'. □

John Lawton is in the Centre for Population Biology and Charles Godfray in the Department of Biology at Imperial College, Silwood Park, Ascot SL5 7PY, UK.

1. Okubo, A., Maini, P.K., Williamson, M.H. & Murray, J.D. *Proc. R. Soc. Lond. B* **238**, 113–125 (1989).
2. Reynolds, J.C. *J. anim. Ecol.* **54**, 149–162 (1985).
3. Shorten, M. *Squirrels* (Collins, London, 1954).
4. Gurnell, J. *The Natural History of Squirrels* (Christopher Helm, London, 1987).
5. Williamson, M.H. & Brown, K.C. *Phil. Trans. R. Soc. Lond. B* **314**, 505–522 (1986).
6. Skellam, J.G. *Biometrika* **38**, 196–218 (1951).
7. Vilet, A.J. thesis, Univ. Oxford (1989).
8. Lubina, J.A. & Levin, S.A. *Am. Nat.* **131**, 526–543 (1988).
9. Mollison, D. *Phil. Trans. R. Soc. Lond. B* **314**, 675–693 (1986).
10. Okubo, A. *Diffusion and Ecological Problems: Mathematical Models* (Springer, Heidelberg, 1980).
11. Lawton, J.H. & Hassell, M.P. *Nature* **289**, 793–795 (1981).

CONTINENTAL PLATES

Flexure reveals great depth

Marcia McNutt

ONE of the outstanding issues in Earth sciences is the question of whether continental and oceanic lithospheric plates are thermally and chemically equivalent beneath the crust, or whether systematic differences in major-element chemistry allow the development of deeper lithospheric roots beneath continental cratons. Resolving this issue is critically important for understanding the long-term growth and stability of continents, the evolution of geochemical reservoirs and mantle dynamics. The analysis of North American gravity data by Bechtel *et al.*, on page 636 of this issue¹, reveals for the first time the variations in the stiffness of the lithosphere across the entire North American continent. The results are most easily interpreted in terms of a lithospheric plate at least 250 km thick beneath the ancient Precambrian Canadian Shield, about twice the thickness that pertains to the more geologically active continental margins and the old ocean basins.

The heterogeneity of continental crust and its long exposure to many tectonic and thermal events have made it difficult to apply the simple thermal-plate model which has been so successful in the ocean basins. For example, with different assumptions concerning the radiogenic heat production of the lower continental crust, heat-flow data can be interpreted as showing either a 125-km-thick plate beneath continents² or one 250 km thick³. From seismic surface-wave velocities and shear-wave travel times, Jordan⁴ has argued that cold lithospheric roots could extend to depths as great as 400 km beneath

the oldest cores of continents.

Another geophysical observation that bears directly on the thermal structure of the lithospheric thermal boundary layer is the thickness of the mechanically rigid elastic layer that supports the weight of anomalous loads placed on, within and beneath the lithosphere. The effective elastic thickness T_e of the lithosphere marks the depth of the transition between elastic and fluid behaviour of rocks subjected to stresses exceeding 100 megapascals over geological timescales.

Because the creep rate of rocks is thermally controlled, in the limit of small amplitude of flexure (such that the finite strength of rocks in the elastic regime is not exceeded), T_e corresponds to the depth of an isotherm, probably somewhere between 600 and 800 °C according to laboratory experiments.

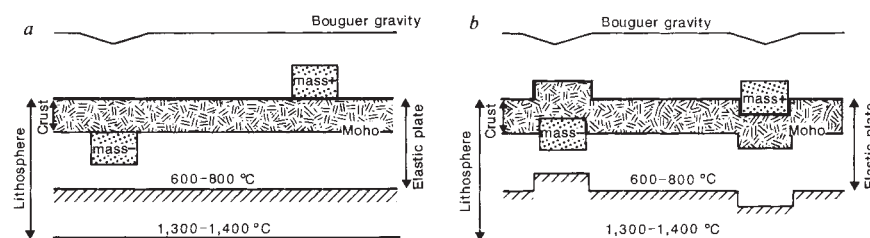
Gravity data are often used to calibrate T_e because the wavelength and amplitude of the gravity anomaly over a load are sensitive to the flexure of the elastic plate induced by that load. To date, gravity data have favoured a thicker thermal boundary layer beneath old continents, as elastic plate thicknesses well in excess of the 40–50-km maximum observed for the oceans have been reported for the stable cores of continents^{5,6}. But those results, primarily obtained from observing the amplitude and wavelength of the gravity anomaly over mountain belts which were superimposed by plate convergence onto stable platforms, are often questioned because many assumptions must be made about the boundary conditions on the elastic plate, the role of subsurface loads and so on.

The map in Fig. 3 of Bechtel *et al.* (see page 638) shows variations in T_e across North America using a relatively new technique based on computing the statistical correlation between Bouguer gravity (the anomaly after correcting for the attraction of all topographic masses) and topography as a function of wavelength in subregions encompassing the entire continent. Their result does not depend on the details of modelling any individual feature, and explicitly allows for subsurface loads (see box below).

According to this analysis, from the young Basin and Range region of the western United States, characterized by

Basis of the coherence technique

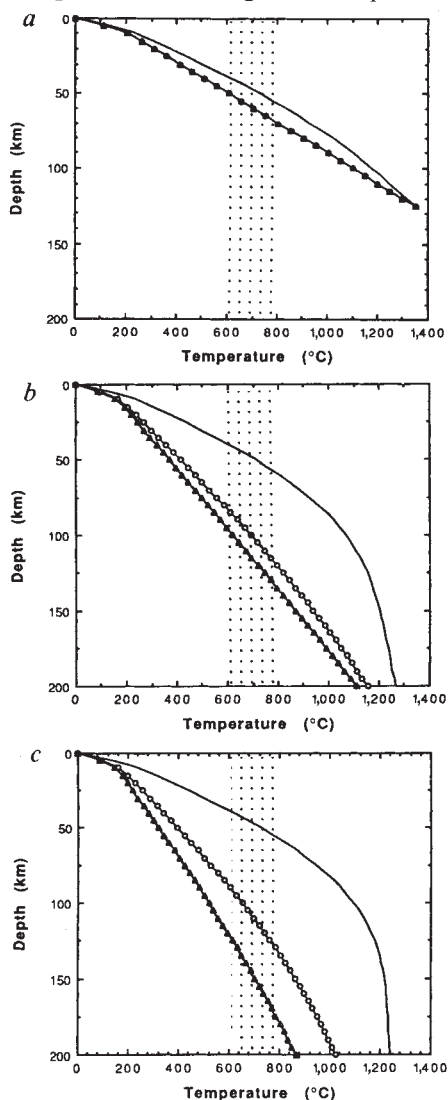
Imagine loads randomly placed on the Earth's surface and at the crust–mantle boundary (the 'Moho'). If the plate is infinitely rigid, *a*, neither surface nor subsurface loads can flex the plate. The topography then measures the magnitude of surface loading, the Bouguer gravity measures the magnitude of subsurface loading, and the lack of correlation between Bouguer gravity and topography at all wavelengths indicates that the plate is perfectly rigid.



If the plate is infinitely weak, *b*, every surface mass will warp the plate to produce a subsurface density anomaly, and every Moho load will lead to surface topography. The correlation between Bouguer gravity and topography will be perfect (and negative) at all wavelengths except very short ones where it approaches zero owing to upward attenuation of the gravity from the Moho loading.

Any realistic case will fall somewhere in between these two extremes, and the particular wavelength at which the correlation changes from –1 at long wavelengths to 0 at short wavelengths is a measure of the stiffness of the plate. □

Tertiary extension and volcanism, to the old, stable Precambrian core of Canada, the thickness of the elastic plate varies from less than 10 km in thickness to values exceeding 100 km. The figure below shows some sample continental geotherms and the predicted thickness of the elastic plate for several values of geological age and lithospheric thickness. Although the T_e values less than 50 km in the geologically active western half of the continent and along the eastern margin are compatible



Geotherms for lithospheric plates 125, 250 and 400 km thick (a, b and c, respectively) computed for lithospheric ages of 100 (—), 500 (○) and 1,000 (△) million years. In all models, the upper continental crust consists of a radiogenic layer 15 km thick with a radioactive heat production of $2.5 \mu\text{W m}^{-3}$. Radioactivity is assumed to be negligible beneath 15 km and the thermal conductivity is set at $2.4 \text{ W m}^{-1} \text{ K}^{-1}$ at all depths. Where each geotherm intersects the shaded region (600–800 °C) is the predicted elastic plate thickness, T_e . For a 125-km-thick plate, T_e will never exceed 70 km, regardless of age. Plates 250 km thick will attain T_e values of 125 km for old continental nuclei older than 1,000 million years. For a 400-km-thick plate, T_e for old continents can be greater than 150 km.

with the 125-km lithospheric plate thickness appropriate for oceanic lithosphere, the simplest interpretation of the large elastic plate thickness over central Canada is that the thermal boundary layer there exceeds 250 km in thickness. (I have not accounted for a possible increase in strength of the continental lithosphere with respect to oceanic lithosphere due to chemical differences, which could be significant if the subcontinental upper mantle is depleted in volatile components which tend to weaken the rocks.)

The fact that the T_e value for old continental lithosphere of North America determined by Bechtel *et al.* is so much larger than the maximum found for old oceanic lithosphere demonstrates that they are not thermally equivalent beneath the crust, which probably also indicates a chemical difference. Geodynamic

DEVELOPMENTAL GENETICS

A long story and a short tail

P. N. Goodfellow

UNTIL the advent of homologous recombination, embryonic stem cells and transgenic mice, genetic analysis of development in mammals relied upon the fortuitous discovery of appropriate mutations. Particularly propitious was the discovery by Dobrovolskaia-Zavadskaja of a mouse with a short tail¹. This animal was a heterozygote for a dominant mutation (*Brachyury*) at the *T* locus. Animals homozygous for the *T* mutation die *in utero* at about day ten of development². From the pattern of embryonic death, it seems that the *T* gene is essential for mesoderm formation and the subsequent development of the notochord^{3,4}. Although of interest in its own right as a developmental mutation, the *T* mutation is better known as the key to the discovery of the *t* complex and its associated effects on development, segregation distortion and fertility^{5,6}. For the past 50 years this combination of features has intrigued geneticists, but it is only in the past decade that the secrets of the *t* complex have yielded to molecular analysis. As reported on pages 617 and 657, Herrmann *et al.*⁷ have returned to the *T* mutation and used a reverse genetic approach to clone the gene, which has allowed Wilkinson *et al.*⁸ to correlate embryonic expression with developmental lethality. It may prove that *Brachyury* is the most interesting developmental mutation associated with the *t* complex.

When *T*-mutant heterozygotes are mated to wild mice up to 50 per cent of the resulting litters contain offspring without tails⁹. If tailless mice from the same litter are intercrossed they produce only tailless mice¹⁰. These crosses were initially interpreted as defining new recessive lethal

arguments suggest that lithospheric thickness greater than 125 km will not be convectively stable unless the density increase from lithospheric cooling is offset by a density reduction from changes in upper-mantle mineralogy, such as via garnet depletion following the extraction of large quantities of basaltic melt⁴. □

Marcia McNutt is at the Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

1. Bechtel, T.D., Forsyth, D.W., Sharpton, V.L. & Grieve, R.A.F. *Nature* **343**, 636–638 (1990).
2. Sclater, J.G., Jaupart, C. & Galson, D. *Rev. Geophys.* **18**, 269–311 (1980).
3. Chapman, C. & Pollack, H. *Nature* **250**, 477–478 (1974).
4. Jordan, T.H. *Rev. Geophys.* **13**, 1–12 (1975).
5. Karner, G.D., Steckler, M.S. & Thorne, J.A. *Nature* **304**, 250–253 (1983).
6. McNutt, M.K., Diamant, M. & Kogan, M.G. *J. geophys. Res.* **93**, 8825–8838 (1988).

mutations, generically called *t* mutants, at the *T* gene. The *t* mutations have no direct effect on tail length but interact with the *T* mutation to produce loss of tails^{10,11}. Another feature of *t* mutations is segregation distortion: male mice carrying a naturally occurring *t* chromosome transmit this chromosome to over 95 per cent of their offspring. The extreme segregation distortion is responsible for the high population frequencies of *t* chromosomes. The *t* mutations could be classified into different complementation groups each of which is characterized by a distinctive pattern of embryonic death⁵. It was proposed that the *t* mutations were defining a complex locus which had a master regulatory role in development⁶. Evidence has been presented that the role of both *T* and *t* genes is exerted through a series of cell-surface molecules expressed on sperm and at crucial moments in development⁶. Unfortunately, the serological results that form the experimental evidence for this hypothesis have proved difficult to confirm¹², and this seductive hypothesis is not consistent with current knowledge of the *t* complex.

The *T* mutation maps to chromosome 17 about 15 centimorgans from H-2. Recombination in this region is severely suppressed between *t* chromosomes and non-mutant chromosomes because of a series of chromosomal inversions shared by *t* chromosomes^{13–15}. Consistent with this observation, recombination occurs freely between different *t* mutant chromosomes and this has allowed a direct genetic analysis of the *t* complex. All the chromosomes carrying *t* mutations are related by descent and carry a series of genes responsible for the segregation distortion,

as well as *tct*, a gene for interaction with *T* to produce loss of tails. Because it is clear that the biology of *t* chromosomes results from a series of genes spread over a large region, the term *t* mutation is not appropriate and most authors refer to the *t* complex. Most theories for the origin of the *t* chromosomes propose that the favourable combination of segregation-distortion genes is fixed by the recombination suppression and that the enormous advantage conferred on the *t* chromosomes by segregation distortion has allowed the accumulation of lethal mutations. The lethal mutations differ between *t* chromosomes and are not allelic^{16,17}. The large number of *t*-lethal mutations is not unusual when it is considered that the *t*-inversions cover at least 15 megabases of DNA (H. Lehrach, personal communication) and even more mutations in this region have been obtained by saturation mutagenesis¹⁸. The *t* complex is an important system for studying segregation distortion and male fertility, but the genetics do not predict that the *t*-lethal mutations are related in any way or that they will, of necessity, define genes important in development. The potential interest of each *t*-lethal mutation must be considered separately.

Notwithstanding the demise of the *t* complex as the universal solution to genetic control of development, *T* is a good candidate for regulating mesoderm formation or interactions. Careful examination of *T*-mutant homozygous embryos indicates that mesoderm cells in the primitive streak behave aberrantly and that the notochord, a derivative of the mesoderm, fails to form correctly. Poor development of the allantois, another mesoderm derivative, is the cause of embryonic death. Shortening of the tail in heterozygotes is also the result of notochord defects stemming from abnormal associations between the notochord and other structures such as the gut and neural tube^{3,4}. Herrmann *et al.*⁷ exploited the existing genetic maps of the *t* region and a large number of cloned random sequences and genes to map sequences in the vicinity of the *T* gene. The sequences were used as probes to construct a long-range restriction map and to define large deletions and duplications known to include the *T* gene. By chromosome walking and jumping, the group isolated new sequences closer to the *T* gene. Finally, a potential HTF island (a CpG-rich region often associated with genes) was identified within a target region of 100 kilobases. The island was associated with a gene, *me75*, which is expressed¹ specifically in mesoderm. Molecular analysis of *T* mutations is consistent with equivalence between *me75* and *T*. Particularly compelling is the analysis of *T*^{Wis}, a new spontaneous mutation of the *T* gene¹⁸. In *T*^{Wis}, the *me75* gene has suffered an insertion of a transposable element.

Nevertheless, the authors rightly point out that final proof of identity between *me75* and *T* will require complementation studies in transgenic mice.

The cloned sequences are a starting point for direct determination of whether the *T* gene is important for regulating the development of mesoderm, or whether it is a bystander required only to meet a specific metabolic challenge. Unfortunately, the sequence of the *me75* gene provides no obvious clue to function, but the predicted protein sequence is not suggestive of a cell-surface molecule. A careful analysis of tissue-specific expression has been more fruitful. As would be expected of the *T* gene, expression of *me75* is restricted to primitive ectoderm (a mesoderm precursor), the mesoderm and the notochord — exactly those tissues known to be affected by mutant *T* genes.

Two fronts can be envisaged for future research. First, the production of antibodies and an investigation of the biochemical properties of the *me75* gene product. Second, genetic investigation of those genes known to modify the effects of the *T* mutations might define proteins that control *T*-gene expression or interact with the *T*-gene product. In addition to *tct*, which may be an allele at the *T* gene¹⁹, there may be a further gene in the *t* region that affects *T* expression²⁰. There is also²¹ a modifier that is unlinked to *T*, as well as numerous mutations known to affect the notochord and tail. Many regard the mouse as the best model for studying human development; it would be ironic if the best developmental system in the mouse was its tail. □

P. N. Goodfellow is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

1. Dobrovolskaia-Zavadskaja, N. C. *r. Séanc. Soc. Biol.* **97**, 114–119, (1927).
2. Gluecksohn-Schoenheimer, S. *Proc. natn. Acad. Sci. U.S.A.* **30**, 134–140 (1944).
3. Chesley, P. J. *exp. Zool.* **70**, 429–459 (1935).
4. Gruneberg, H. J. *Embryol. exp. Morph.* **6**, 424–443 (1958).
5. Gluecksohn-Waelsch, S. & Erickson, R.P. *Curr. Top. dev. Biol.* **5**, 281–316 (1970).
6. Bennett, D. *Cell* **6**, 441–454 (1975).
7. Herrmann, B.G., Labeit, S., Poustka, A., King, T.R. & Lehrach, H. *Nature* **343**, 617–622 (1990).
8. Wilkinson, D.G., Bhatt, S. & Herrmann, B.G. *Nature* **343**, 657–659 (1990).
9. Dunn, L.C. *Science* **144**, 260–263 (1964).
10. Dunn, L.C. *Proc. natn. Acad. Sci. U.S.A.* **23**, 474 (1937).
11. Dobrovolskaia-Zavadskaja, N. & Kobozieff, N. C. *r. Séanc. Soc. Biol.* **110**, 782–784 (1932).
12. Goodfellow, P.N., Levinson, J.R., Gable, R.J. & McDevitt, H.O. *J. reprod. Immunol.* **1**, 11–21 (1979).
13. Silver, L.M. & Artzt, K. *Nature* **290**, 68–70 (1981).
14. Herrmann, B.G. *et al. Cell* **44**, 469–476 (1986).
15. Hammer, M.F., Schiemann, J. & Silver, L.M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3261–3265 (1989).
16. Artzt, K., McCormick, P. & Bennett, D. *Cell* **28**, 463–470 (1982).
17. Condamine, H., Guenet, J.-L. & Jacob, F. *Genet. Res.* **42**, 335–344 (1983).
18. Sheddlovsky A., King, T.R. & Dove, W.F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 180–185 (1988).
19. Justice, M. & Bode, V. *Mouse News. Lett.* **80**, 168–169 (1988).
20. Nadeau, J.H., Varnum, D. & Burkart, D. *Genetics* **122**, 895–903 (1989).
21. Artzt, K., Cockingham, J. & Bennett, D. *Genetics* **115**, 601–605 (1987).

Sweating it out

A FRESHLY-MADE bed feels cool, not merely because its sheets are at room temperature, but because they are dry. Most bed-clothes are hygroscopic and absorb water vapour from the body, which is cooled accordingly. Once they have taken up all they can hold, the bed warms up. When the bed is aired, the water evaporates again.

So Daedalus is inventing bed-clothes and garments with controllable humidity. Intrepid DREADCO volunteers are wrestling with sheets loaded with silica gel, or donning underwear saturated with hydrated sodium carbonate, to judge the cooling or warming effect of a controlled dry or damp personal environment. Meanwhile, DREADCO's physicists are developing special water-pumping garments, by which the wearer can alter his personal humidity at will.

Their cunning new 'Watershed' fabric has porous, metallized electrode surfaces on each side. A voltage applied across the fabric impels its water content towards the positive side by electro-osmosis. This side therefore evolves water vapour, whereas the negative side absorbs it. Set to pump water vapour outwards, DREADCO's Watershed clothing will cool the wearer; set to pump it inwards, it will envelop him in a warming internal 'fog'.

To pump effectively against a reasonable vapour-pressure gradient, Watershed will need an electro-osmotic potential of several kilovolts. At first Daedalus feared that this would electrocute the wearer. But, in fact, a Watershed garment needs to pump so little water per unit area that it will only draw a tiny, almost electrostatic current. Its high voltage will then be hardly more inconvenient than that of a hastily donned nylon shirt. Indeed, Daedalus hopes to generate the required voltage from the wearer's own movements, through frictional charge-separation units in rubbed areas of the garments like sleeves or trouser legs. The wearer thus acts as his own thermostat. The more vigorously he exercises, the more strongly his Watershed clothing will cool him. In the cold he will reverse the polarity, when the more he shivers the faster it will warm him up. He will have full personal, automatic air-conditioning.

Profound social consequences will follow. The massive power drain of heating and air-conditioning will plummet: why cool a whole building when the inhabitants can cool themselves? Fashion designers will relish the new metallized high-tech fabric, until they realize with horror that it spells the end of fashion. For one single, universal Watershed overall will serve every need — study or vigorous exercise, indoors or outdoors or even in the wet. It can even pump out the rain to keep you dry and smart in the most drenching downpour. David Jones

The dissolution of semantics

SIR—In a previous report¹, we described a patient with an impairment of comprehension apparently restricted to certain semantic categories in the auditory-verbal domain. This category and modality specific disorder was associated with a local degenerative condition². We now describe the further degeneration of semantic knowledge we have observed.

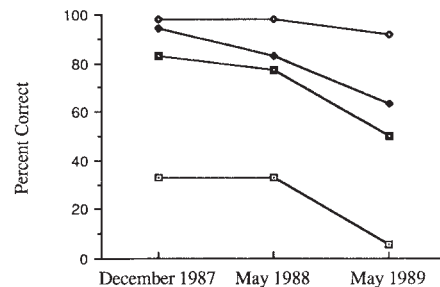
When first assessed, the patient (T.O.B.) had a consistent impairment of his ability to comprehend the spoken names of living things, but remarkably, he had no difficulty in comprehending pictures of the same items. His comprehension of objects was good with both visual and auditory presentation. Because his impairment was specific both to the modality of presentation and to a semantic category, we challenged the orthodox view that knowledge is represented in a "general purpose" semantic system. Our conclusion, neatly summarized by Marshall³, was that "semantic information is multiply represented in the normal brain and this duplication is linked to the input modalities whereby we gain knowledge of the world".

T.O.B. has since been seen regularly for clinical counselling and for the serial assessment of his comprehension difficulties. His condition remained remarkably stable for six months (December 1987 to May 1988), but by May 1989 there had been a clear deterioration in his verbal comprehension. He remained able to express himself volubly, but became increasingly circumlocutory and reliant on stereotyped phrases.

In the original test session, T.O.B. was asked to define the spoken names of 48 objects and 48 living things, as well as to 'define' pictures of the same stimulus items. Although he was unable to name the picture stimuli, he was often able to provide adequate 'picture definitions' conveying the core concept. Thus, when asked to define the spoken word 'dolphin', T.O.B. replied "a fish or a bird", whereas when shown a picture of a dolphin he responded "lives in water . . . they are trained to jump up and come out . . . In America during the war years they started to get this particular animal to go through to look into ships". He was subsequently re-tested on this same stimulus pool 6 and 18 months later (see figure). The results show a marked decline in T.O.B.'s comprehension, but even on his most recent assessment he still shows effects of modality of presentation and semantic category.

T.O.B.'s knowledge of the spoken names of living things has shown a significant decline (binomial test $N=14$, $x=0$, $P<0.001$); his knowledge of pictures of living things (which was previously rela-

tively intact), is now significantly worse (binomial test $N=15$, $x=0$, $P<0.001$). In the case of objects, there has been a deterioration in his knowledge of many spoken names that he knew previously (binomial test $N=16$, $x=0$, $P<0.001$). For example, in 1987 he was able to define the word



T.O.B. Percentage correct on three successive assessments. Open squares, living things (auditory); filled diamonds, living things (visual); filled squares, objects (auditory); open diamonds, objects (visual).

'canoe' as a "small version of a boat, long and narrow, one or 2 people" whereas, by 1989, all he could say was "a common word to me". But there has been no significant deterioration for one category of knowledge, namely visual objects. For example, T.O.B. was able to describe a picture of a thermometer as "for temperature, automatically records the actual temperatures that there are". Indeed, were one to have examined T.O.B. for the first time at this stage in his illness he might well have been characterized as a

case with an 'island' of preserved knowledge specific to inanimate objects in the visual modality. The pattern of impairment on the most recent serial assessment is the obverse of that documented on the first.

Because the selective preservation of knowledge may be as informative as its selective loss, we would argue that the present evidence adds further weight to the contention that there are modality and category specific cerebral meaning systems. This conclusion is strengthened by the documentation of different rates of degeneration in T.O.B.'s verbal and visual knowledge. Were a single 'semantic' store implicated, it would be hard to explain why verbal knowledge should deteriorate more rapidly than visual knowledge, and still less why a category-specific profile of impairment should be maintained.

ROSALEEN MCCARTHY*

Department of Experimental Psychology,
University of Cambridge,
Downing Street,
Cambridge CB2 3EB,
UK

ELIZABETH K. WARRINGTON

Department of Psychology*,
The National Hospitals,
Queen Square,
London WC1N 3BG,
UK

1. McCarthy, R.A. & Warrington, E.K. *Nature* **334**, 428-430 (1988).

2. Tyrell, P.J., Warrington, E.K., Frackowiak, R.S.J. & Rossor, M.N. *Brain* (in the press).

3. Marshall, J.C. *Nature* **334**, 378 (1988).

Genetics of long-spurred pheasants

SIR—von Schantz *et al.* recently reported¹ that long-spurred male pheasants (*Phasianus colchicus*) survive better than short-spurred males, are preferred as mates by female pheasants, and that those females mating with long-spurred males produce more offspring. These results are of great significance to current controversies in sexual selection theory^{2,3}. But the data of von Schantz *et al.* in fact challenge their interpretation that the genetic superiority of long-spurred males is directly responsible for the increased reproductive success of females mating with them.

von Schantz *et al.* quantified female reproductive success by the number of chicks hatched. In most birds, the most important cause of variation in hatching success is predation⁴. Because differences in the male genetic contribution to eggs should not affect their vulnerability to predators, the effects of predation should be removed from the analysis before assessing whether paternal quality directly influences hatching success. Because predation is usually absolute (that is, all eggs in a clutch are destroyed), the effect

of predation can be largely eliminated by excluding from the analysis all clutches in which no eggs hatched. (In nests suffering no predation, hatching success usually exceeds 80%, see ref. 5.) If clutches hatching no young are eliminated from Fig. 1b of ref. 1, the significant correlation between male spur length and female reproductive success disappears ($r^2 = 0.02$, $P = 0.45$, $N = 30$).

What this implies is that the increased reproductive success of females mating with long-spurred males arises, not because they lay eggs of intrinsically high viability, but because their nests suffer less predation. As males do not contribute to the care of nests¹, reduced predation cannot be attributed to better parenthood by long-spurred males. Perhaps, as Kirkpatrick² suggested in his *News and Views* article accompanying ref. 1, female pheasants are behaving like Burley's⁶ female zebra finches, providing better care for eggs sired by superior males. It remains to be demonstrated that females mating with superior males have higher reproductive success independent of their

own reproductive effort.

PATRICK J. WEATHERHEAD
DREW J. HOYSAK

Department of Biology,
Carleton University,
Ottawa, Ontario K1S 5B6,
Canada

VON SCHANTZ *ET AL.* REPLY—Weatherhead and Hoysak are correct that the significant correlation between male spur length and female reproductive success disappears when females hatching no chicks are removed. But we are not convinced that their reanalysis is relevant. First, we find no basis for their challenge of our interpretation that “the genetic superiority of long-spurred males is directly responsible for the increased reproductive success of females mating with them”. Data on the survival rates of chicks in relation to their paternity would have been necessary to reach such a conclusion, but were not available. Rather, we stated explicitly (p.168 of ref. 1) that “We do not know to what extent male genetic quality affects the number of hatched chicks and offspring survival”, which is clearly divergent from the interpretation Weatherhead and Hoysak force upon us.

Of the 15 females hatching no chicks (Fig. 1b in ref. 1), only three suffered from nest predation. A removal of the effect of nest predation, as Weatherhead and Hoysak suggest, does therefore not affect the correlation substantially ($r=0.337$, $P<0.02$, $N=42$) so that it is not correct to conclude that predation was the most important cause of variation in female reproductive success.

The remaining “unsuccessful” females either did not attempt nesting ($N=9$) or abandoned their intact nests ($N=3$). The spur length was significantly shorter among the mates of the nine non-nesting females than among the mates of females that attempted nesting at least once ($t=2.51$, $P<0.02$, $N=45$). Indeed, these data support Burley’s hypothesis⁶ that mates of attractive individuals invest more in reproduction. But it remains a controversial question whether the relationship between female reproductive success and mate spur length arises from females’ “differential allocation”, assuming inheritance of attractiveness⁶, the “good genes”, assuming inheritance of genotypic quality^{7–10}, or in some other way. Accordingly, we hesitated to speculate prematurely on

the causes of our observations.

The two main alternatives are, in any case, by no means mutually exclusive; if maternal effort affects female reproductive success, as may well be the case, the two hypotheses are difficult to separate. Although we are well aware of Burley’s hypothesis and fully accept Weatherhead and Hoysak’s point, it is in our opinion crucial that our data are not relevant either to the refutation or the promotion of either of the two overlapping hypotheses. In the case of the pheasants we have studied, both processes lead to the same effect, that is sexual selection for a viability-based male trait¹.

TORBJÖRN VON SCHANTZ
GÖRGEN GÖRANSSON
GUNILLA ANDERSSON
MATS GRAHN
HÅKAN WITZELL*

Departments of Animal Ecology and
Theoretical Ecology*,
Ecology Building,
University of Lund,
S-223 62 Lund, Sweden

Synthetic tips

SIR—Capabilities in scanning tunnelling and atomic force microscopy (STM and AFM) depend on probe-tip properties, but tip properties are at present poorly controlled; suitable tips have poor reproducibility at the atomic level. Techniques for precise control of the mechanical and chemical nature of tip structures would be of evident value. We therefore propose a ‘molecular tip’ made of a protein-like molecule which has been designed to bind to a crystal-corner STM/AFM tip, and which could potentially be tailored, manufactured and reproducible.

Advances in protein engineering¹ and synthesis of novel molecular receptors² demonstrate a growing capability for designing large molecules, including those that bind to other molecular structures. We note that crystal corners can have atomically regular equilibrium structures with radii in the nanometer range³, a size compatible with such binding. If a combination of tip and bound tip-capping molecule could be developed, this could open a wide range of experimental possibilities.

Given a suitable tip-capping molecule, techniques for molecular design and synthesis could be exploited to refashion the working surface opposite the tip-binding surface. Diverse properties should be achievable. In particular, while STM technology has enabled chemical modification of single molecules⁴, greater specificity is a significant goal⁵. One approach would exploit positional control of interchangeable reactive moieties on scanning probe microscope tips to favour chemical reactions at specific sites on

product molecules bound to a substrate.

Consider an AFM tip capped by a protein (-like) molecule with a working surface incorporating a binding site for a ligand, which in turn bears a reactive moiety. AFMs with tips characterized by a small (~ 1 N/m) vertical spring constant⁶ and a large ($\geq 10^3$ N m⁻¹) transverse spring constant (T. Albrecht, personal communication) can be positioned with atomic precision in a liquid medium⁶. Since the spring constants for deforming multi-nanometer scale globular proteins and for bending single bonds are both >10 N m⁻¹, there seems no obstacle to positioning bound reactive moieties with a transverse spring constant $\geq$ N m⁻¹.

Ligands of similar structure can bear differing reactive moieties; these may be interchanged (given suitable binding-site properties) by adjusting ligand solution concentrations. A 10^{-4} M concentration would yield high tip occupancy given a dissociation constant $\leq 10^{-5}$ M, and interchange on a 0.1 s time scale given rate constants $\geq 10^5$ M⁻¹ s⁻¹ for association and ≥ 10 s⁻¹ for dissociation. For protein-ligand interactions, rate constants $\geq 10^8$ M⁻¹ s⁻¹ and ≥ 100 s⁻¹ are compatible⁷.

Positioning of reactive moieties by proteins can yield high effective concentrations, $>10^2$ M even for mobile surface-residue thiol groups⁷. With a bound-ligand effective concentration at the tip $\geq 10^2$ M and a $\leq 10^{-4}$ M concentration in solution, an AFM-based mechanism could create a positionable, site-specific concentration enhancement $\geq 10^6$. At 300 K, a thermally excited structure with a 1 N m⁻¹ spring constant will be characterized by a gaussian positional uncertainty with sigma <0.065 nm; the ratio of effective concentration enhancement between a target site and a site 0.25 nm distant will then be $>10^3$. Thus, strong and adjustable positional control of effective concentrations (and hence of reaction rates) on an atomic distance scale appears physically possible in a class of systems that may be practically realizable.

K. ERIC DREXLER

Department of Computer Science,
Stanford University,
Stanford, California 94305,
USA

JOHN S. FOSTER

IBM Research Division,
Almaden Research Center,
San Jose,
California 95120,
USA

1. von Schantz, T. *et al.* *Nature* **337**, 166–169 (1989).
2. Kirkpatrick, M. *Nature* **337**, 116–117 (1989).
3. Pomiankowski, A. *Nature* **337**, 696 (1989).
4. Ricklefs, R.E. *Smithsonian Contrib. Zool.* **9**, 1–48 (1969).
5. Johnsgard, P. *The Pheasants of the World* (Oxford University Press, 1986).
6. Burley, N. *Am. Nat.* **132**, 611–628 (1988).
7. Zahavi, A. *J. theor. Biol.* **67**, 603–605 (1977).
8. Hamilton, W.D. & Zuk, M. *Science* **218**, 384–387 (1982).
9. Kodric-Brown, A. & Brown, J.H. *Am. Nat.* **124**, 309–323 (1984).
10. Andersson, M. *Evolution* **40**, 804–816 (1986).

1. DeGrado, W.F., Wasserman, Z.R. & Lear, J.D. *Science* **243**, 622–628 (1989).
2. Lehn, J.-M. *Angew. Chem. Int. Ed. Engl.* **27**, 89–112 (1988).
3. Fink, H.W. *IBM J. Res. Dev.* **30**, 460–465 (1986).
4. Foster, J.S., Frommer, J.E. & Arnett, P.C. *Nature* **331**, 324–326 (1988).
5. Drexler, K.E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5275–5278 (1981).
6. Drake, B. *et al.* *Science* **243**, 1586–1589 (1989).
7. Creighton, T.E. *Proteins* 1–515 (Freeman, New York, 1984).

Fracture without fusion

SIR—Several years ago, Derjaguin *et al.*¹ claimed that the fracture of D₂O by macroscopic projectiles with moderate velocities induces d + d fusion. The evidence, scant at best, was based on detecting pulses in a BF₃ proportional counter array, in coincidence with the macroscopic impacts, in excess of the numbers observed using H₂O targets. More recent work by the same group claims that fusion also occurs in mechanically agitated mixtures of titanium in the presence of deuterium². These intriguing claims gain some credibility from studies of exoelectrons^{3,4} which indicate that a significant fraction of the electrons emitted on the fracture of certain materials have energies exceeding 1 keV. Thus, Dickinson *et al.*⁴ have found that approximately 15 per cent of the electrons

with large projectiles are limited, we have repeated this experiment. Our results have been negative, with an upper limit of the number of neutrons per fracture event which is 15-times less than the positive result of Derjaguin *et al.*¹.

In our experiment, a gas gun propelled steel ball bearings, of mass 0.131 g and diameter 1/8 inch, at $168 \pm 4 \text{ m s}^{-1}$ at polycrystalline ice. The ice (the D₂O was 99.9 atom per cent D) was frozen in a dry-ice acetone bath at -66°C . The target was inside the well of a liquid scintillator detector (LS). Neutron pulses in the liquid scintillator were selected by standard 1-dimensional pulse shape discrimination logic. A set of twelve proportional counters (PCs), imbedded in a paraffin cylinder, encircled the target and the liquid scintillator. Pulses from the PCs result from the capture of thermalized neutrons, whereas the pulses from the scintillator result from neutron scattering. Another set of three PCs positioned at a greater radial distance, outside both Cd and B₄C neutron shielding, was used as a guard ring, run in anticoincidence with the main PC ring. The two detector systems (LS and PC) were operated both in singles modes and in coincidence with each other. The efficiencies, background count rates and the upper limits (2 σ) to the number of neutrons per shot for the three modes of

operation are given in the table.

The results from 75 shots on both D₂O and H₂O are shown in the figure. There is no indication of enhanced count rate at the time of the shot (time=0) or afterwards. Analysis of the spectra yields an upper limit less than 1 neutron per shot

Neutron counter characteristics and results

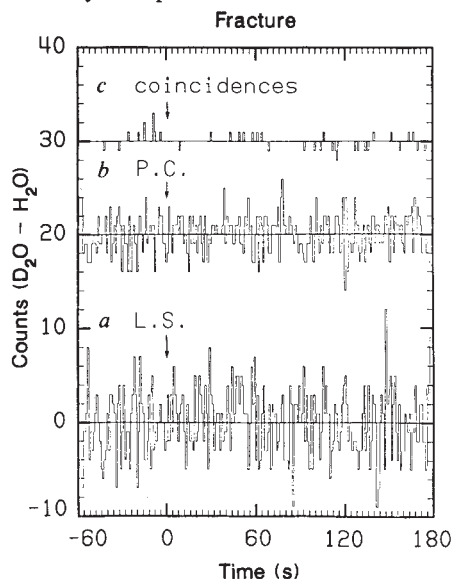
	LS	PC	LS-PC
Background (counts min ⁻¹)	5.3	1.6	0.054
Efficiencies with neutrons from ²⁵² Cf (2 ± 1 MeV)	25(%)	13.8(%)	4(%)
Am-Be (4 ± 2 MeV)	~20(%)	11.6(%)	4(%)
Upper limit neutrons per shot	0.6	0.8	0.7

(see Table). The work reported in ref. 1 gives a positive result at a level of approximately 10 neutrons per shot, yet as best as we can determine, our experimental conditions are similar in all important details. The limited statistics presented in refs 1 and 2 and our negative results convince us that there is no compelling evidence for fracture induced fusion.

L. G. SOBOTKA
P. WINTER

Departments of Chemistry and Physics,
Washington University,
St Louis,
Missouri 63130,
USA

1. Derjaguin, B.V., Kluev, V.A., Lipson, A.G. & Toporov, Yu. P. *Coll. Jour. USSR* **48**, 8–10 (1986).
2. Derjaguin, B.V., Lipson, A.G., Kluev, V.A., Sakov, D.M. & Toporov, Yu. P. *Nature* **341**, 492 (1989).
3. Vladikina, T.N., Derjaguin, B.V., Kluev, V.A., Toporov, Yu. P. & Khrustalev, Yu. A. *Trans. ASME* **102**, 552–555 (1980).
4. Dickinson, J.T., Donaldson, E.E. & Park, M.K. *J. Mat. Sci.* **16**, 2897–2908 (1981).
5. Beuhler, R.J., Friedlander, G. & Friedman, L. *Phys. Rev. Lett.* **63**, 1292–1295 (1989).



The difference spectra for 75 shots on D₂O and 75 shots on H₂O. *a*, Singles liquid scintillator; *b*, singles proportional counter array; *c*, coincidence between the liquid scintillator and the proportional counter array. The time of the shots is $t=0$. *b* and *c* are plotted with offsets of 20 and 30 counts, respectively.

emitted on the fracture of epoxy strands could not be stopped by retarding potentials in the 500–1,000 V range.

The recent work of Beuhler *et al.*⁵ on cluster impact fusion may also relate to this problem. Fusion cross-sections, from the impact of deuterated clusters of ice containing a few hundred molecules on targets with a high deuteron density, are substantially greater than would be expected from gas phase reactions with similar nucleus–nucleus velocities.

The works cited above suggest that there is a need for further work to investigate the possible fusion mechanisms which might occur in the fracture of solid targets with high deuteron densities. Because the statistics of the measurements of the shattering of deuterated ice

Folding pathway enigma

SIR—Fersht and co-workers recently published an interesting approach to studying the transition state and pathway of protein unfolding¹. They propose a new way to interpret the kinetics of urea-induced unfolding of wild-type and mutant forms of barnase. They find significant differences in the importance of specific protein regions on the energy level of the transition state for unfolding. Their interpretation is that some regions are already exposed to solvent whereas others, such as the N-terminal helix, still show native-like structure in the transition state.

These experiments provide additional insight into the mechanisms of urea-induced protein unfolding, but they do not necessarily provide information on the transition state and pathway of refolding of proteins as suggested by the authors.

Their data analysis is based on the assumption that barnase exists only in two distinct conformational states: native and denatured. The validity of such a two-state model could only be proven by

calorimetric measurements which, to our knowledge, have not yet been carried out.

Also, the thermodynamic cycles proposed in the article (see Fig. 4, ref. 1) are purely hypothetical, as the authors have to assume an equilibrium between wild-type and mutant protein which is obviously out of range of experimental handling.

The main shortcoming, however, is the use of the unfolding kinetic data to elucidate the transition state and pathway of protein folding. This is only legitimate under identical conditions, where the principle of microscopic reversibility is valid, that is, under strongly denaturing conditions. The authors' comparison of their results with recent NMR studies on early refolding steps^{2,3} is misleading because these studies were carried out under "strongly native conditions". Under these conditions (as Fersht and co-workers agree) the folding pathway is essentially different from the mechanism of unfolding because partially folded intermediates are formed rapidly in the refolding of many proteins^{4–6}.

For these reasons, it is unlikely that the unfolding approach will provide new insight into the still enigmatic process of protein folding.

JOHANNES BUCHNER

*Institut für Biophysik und Physikalische Biochemie,
Universität Regensburg,
Universitätsstrasse 31,
D-8400 Regensburg,
FRG*

THOMAS KIEFHABER

*Laboratorium für Biochemie,
Universität Bayreuth,
Postfach 10 12 51,
D-8580 Bayreuth,
FRG*

FERSHT *ET AL.* REPLY—Any novel analysis will, and should, attract intense scrutiny. All of the points that have been raised have precedents, but we welcome the opportunity to discuss them. The analysis of a forward pathway by studying the reverse, for example, is well-established in physical chemistry and enzymology¹ and in protein folding².

First, the use of hypothetical thermodynamic cycles is perfectly valid and is frequently employed in thermodynamic analysis; for example, the classic Born-Haber cycle described in most elementary physical chemistry texts. This is allowed by the first law of thermodynamics, which states that energy changes depend only on initial and final states and not on the pathway between them. Cycles equivalent to those we use³ are also used by the theoreticians for their free-energy calculations (see, for example, ref. 9).

Second, the data analysis is not based on the assumption that barnase exists in only two distinct conformational states. For the equilibrium unfolding cycle, the analysis is based simply on the energy difference between initial (folded) and final (unfolded) states and the first law of thermodynamics. Just as the law allows us to add hypothetical intermediates, it allows real ones to be ignored. For the kinetic cycle, as explained in the article¹, transition-state theory when applied to the unfolding kinetics avoids the problems of intermediates accumulating, because this happens before the rate-determining step only on the refolding pathway. Incidentally, microcalorimetry is not the only way of detecting folding intermediates. Kinetics can detect intermediates that are not observed by equilibrium microcalorimetry and, from measurements of both folding and unfolding rates, either show that a two-state transition is obeyed or that an intermediate accumulates (ref. 10, and our unpublished work). We did indeed, detect an intermediate on the refolding pathway of barnase by kinetic methods and H-D exchange (see refs 2, 3), but its presence

does not affect our analysis of unfolding as it occurs after the rate-determining step in the unfolding direction and accumulates only slightly in the transition region (unpublished data).

Third, we stated clearly that microscopic reversibility applies under identical reaction conditions and for a reversible process (ref. 1; see also ref. 7, p. 89–90). It holds not just for “strongly denaturing conditions” but for all concentrations of urea, because folded and unfolded protein always exist in equilibrium even though the equilibrium constant may be far from unity. At any one of these concentrations, the transition states for folding and unfolding are the same. The question is whether or not the transition state changes with change of denaturant. Creighton¹¹ has summarized the evidence that the nature of the transition state for folding–unfolding does not change with denaturant concentration—the rate of unfolding changes uniformly with changing conditions but refolding changes non-uniformly because the nature of the unfolded state changes⁴. (This is a principal reason why unfolding kinetics is easier to analyse.)

This laboratory has espoused the need for determining the full free-energy pathways of reactions by measuring rate constants in both the forward and reverse directions (see studies on the tyrosyl-tRNA synthetase). We have measured the rate constants for refolding of barnase and many of its mutants and combined these with the unfolding data to characterize the structure of the folding intermediate and the transition states by extension of our protein engineering analysis (manuscript in preparation). The combined data give new information on the folding intermediate but no more information on the transition state than did the unfolding data alone.

Creighton has stated: “Instead of searching for nucleation sites in unfolded proteins, it might be more relevant to search for unfolding nucleation events in the native conformation” (ref. 11). For this reason, and for the necessity of constructing free-energy profiles, it is an essential element of folding studies to measure unfolding rate data. Furthermore, because refolding kinetics is fraught with complications arising from the presence of folding intermediates, *cis-trans* isomerization of prolines, and chemical processes that occur when the denatured protein is incubated for many minutes, unfolding kinetics is simpler to analyse and far less prone to artefact. Unfolding studies by themselves can be used to define the last transition state on the refolding pathway. Unfolding studies might also be more relevant than refolding for relating to phenomena *in vivo*, because the pathway of folding *in vivo* depends on the sequential formation of proteins during translation and may be aided by

chaperones, whereas unfolding does not depend on these factors. Far from the unfolding approach being unlikely to provide new insight into protein folding, it should be clear from the above that unfolding studies are an essential element in the understanding of folding, and are just as important as refolding studies.

A.R. FERSHT

J.T. KELLIS JR

A.T.E.L. MATOUSCHEK

L. SERRANO

*MRC Unit for Protein Function and Design,
University Chemical Laboratory,
Lensfield Road,
Cambridge CB2 1EW,
UK*

1. Matouschek, M., Kellis, J.T., Serrano, L. & Fersht, A.R. *Nature* **340**, 122–126 (1989).
2. Udgoankar, J.B. & Baldwin, R.L. *Nature* **335**, 694–699 (1988).
3. Roder, H., Elove, G.A. & Englander, S.W. *Nature* **335**, 700–704 (1988).
4. Schmid, F.X. & Baldwin, R.L. *J. molec. Biol.* **135**, 199–215 (1979).
5. Kim, P.S. & Baldwin, R.L. *A. Rev. Biochem.* **51**, 459–489 (1982).
6. Kuwajima, K. *Proteins* **6**, 87–103 (1989).
7. Fersht, A.R. *Enzyme Structure and Mechanism*, 89–90, 205–206, 237, 427 (Freeman, New York, 1985).
8. Segawa, S.-I. & Sugihara, M. *Biopolymers* **23**, 2475–2488, 2489–2498 (1984).
9. Rao, S.N., Singh, U.C., Bash, P.A. & Kollman, P.A. *Nature* **328**, 551–554 (1987).
10. Kuwajima, K., Mitani, M. & Sugai, S. *J. molec. Biol.* **206**, 547–561 (1989).
11. Creighton, T.E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5082–5086 (1988).

Who's who?

SIR—The Haldane beetle story alluded to by Reg Passmore¹ and previous writers belongs to J. B. S., who made the remark in 1951 in a lecture on the biological problems of space flight to the British Interplanetary Society, whose journal then reported it² (as discussed in my book *Comparative Social Recognition*³).

PATRICK COLGAN

*Department of Biology,
Queen's University,
Kingston,
Canada,
K7L 3N6*

SIR—The Haldane beetle story⁴ was an invention by Hutchinson to tease Haldane; J. B. S. Haldane was the Haldane referred to and the story has no connection whatever with either Jowett or Oxford. The apocryphal conversation would have taken place, if it had happened, in a common room in University College, London.

MARK WILLIAMSON

*Department of Biology,
University of York,
York YO1 5DD,
UK*

1. Passmore, R. *Nature* **342**, 866 (1989).
2. Slater, A.E. *J. Br. Interplanet. Soc.* **10**, 154–158 (1951).
3. Colgan, P. *Comparative Social Recognition*. P.10 (Wiley, New York, 1983).
4. Hutchinson, G.E. *Am. Nat.* **93**, 145 (1959).

Ways of thinking

Edward S. Reed

The Remembered Present: A Biological Theory of Consciousness. By Gerald Edelman. Basic Books: New York, 1990. Pp.384. \$29.95.

In his unfinished *Project for a Scientific Psychology*, Freud claimed that local neuronal activity constitutes the units of our conscious life. Like Freud, many neurophysiologists and philosophers have assumed that a sufficiently intelligent neuroscientist equipped with on-line brain-scanning technology could, in some sense, 'read off' an individual's mental state from their brain states. Freud soon abandoned his thesis and adopted a form of dualism to explain consciousness. This limited choice, between a theory of 'gnostic units' (such as 'grandmother' cells) and some kind of dualism, has been the unhappy lot of those who have attempted to relate brain and consciousness in the past century.

In *The Remembered Present*, Gerald Edelman attacks the now dominant view that mental states are made up of neural units, but he does so without retreating into dualism, and with rigorous attention to biological detail. He argues that the gnostic-unit theory is an example of what immunologists would criticize as 'instructionist thinking', the idea that the environment 'instructs' particular neurons (either directly, through experience, or indirectly, via genes) to perform a particular psychological function. For Edelman and a growing number of neuroscientists, 'selectionism' offers an attractive alternative to instructionism. Selectionist theories are based on the idea that heredity supplies the nervous system with an overabundance of differentially functioning neurons, together with a set of epigenetic constraints such that experience creates populations of neurons (neuronal groups) which accomplish psychological functions with differential rates of success. What any individual neuron does is thus in part a function of its connections, and these change with experience. Edelman's two earlier books, *Neural Darwinism* and *Topobiology*, provide a richly detailed theory of neuronal group selection, a theory well summarized in chapter 3 of the present volume.

The most important implication of selectionism for a neural theory of mental activity is elaborated in chapter 4. If only neuronal groups can be said to have any specific psychological functions, then a theory of mental life needs to be concerned with the role of changing interconnections within and among these groups of neurons, not with individual neurons. Chapters 6 to 13 develop this insight in what can only be called breath-

taking scope and considerable detail, systematically relating neuronal processes with the components of William James's "stream of consciousness".

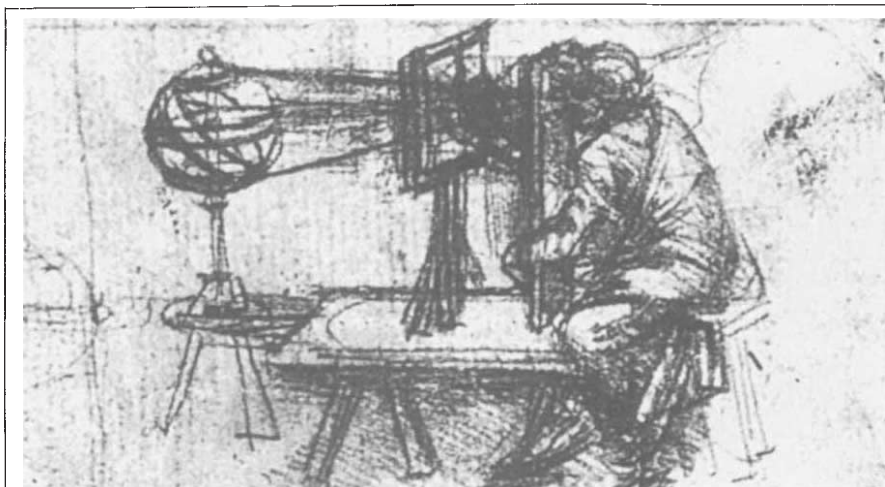
The heart of Edelman's theory is a series of arguments concerning how connections among neurons allow neuronal groups to integrate and cross-correlate their processes, what Edelman calls "re-entrant" connections among neuronal groups. All the models begin with the fundamental insight that the mere detection of a feature of stimulation (no matter how precise) is insufficient to yield knowledge of specific environmental events. The detection of these events (and the corresponding control of action) requires for its accomplishment a "classification couple or n -tuple" of interconnected groups. In such a network there will be first, several groups of different kinds of neural feature detectors, with crosslinkages among them; and second, neuronal groups specifically sensitive to covariation in the activity of the first kind of groups. The models of visual processing presented in chapter 5 which make use of this principle are among the most significant results yet for any neural networks. They show how interconnections among highly specialized groups can be integrated to yield flexible and accurate perception at the object level.

To move from detection of environmental realities to abstraction or conception requires what Edelman calls a "global mapping" that cross-correlates the performance of populations of n -tuples (chapters 4 and 5). But abstraction is not a

once-and-for-all process. New information often requires changes in classifications, and new needs of the animal can also cause reorganization of concepts. Hence, there is a continual process of "recategorization" which involves re-entrant connections in global mappings.

Edelman's argument here is in the tradition of the psychologist Sir Fredric Bartlett who more than 50 years ago emphasized that memory is a process of ongoing conceptualization, not mere storage of ideas. The arguments in chapter 6 concerning memory as recategorization, and the subsequent discussions of the role of memory in self-awareness in chapters 9, 11 and 13 are an outstanding contribution to neuroscience. Heretofore neurophysiologists have not taken seriously the kinds of evidence and arguments, long known to psychologists, which undermine the 'storehouse' metaphor of memory. Hence the peculiar situation has arisen in which most neuroscientists working on memory have been looking for neural underpinnings of phenomena which are very likely to be non-existent!

If a creature is to be conscious enough to maintain a systematic course of action in the face of the complexities of any habitat, it must be capable of selecting priorities: what to look for, what to act on, in what sequence, and with what contingencies. To account for this ordering of concepts, Edelman develops a remarkable series of models concerning what he calls the "cortical appendages", basal ganglia, cerebellum and hippocampus (chapters 7 and 8). He shows in considerable detail how re-entrant connections between cortex and these extracortical "organs of succession" could enable the nervous system to select, order and organize its global mappings. The function of the hippocampus is to integrate re-entrance over time periods much longer than perceptual sampling times, allowing dynamic integration of interoceptive



In perspective — Leonardo da Vinci's *Draughtsman using a transparent plane to draw an armillary sphere* (c.1510) is reproduced from *The Science of Art*, in which Martin Kemp discusses optical themes in Western art. Published by Yale University Press, £45, \$60.

and exteroceptive processes. The basal ganglia modulate inter-cortical relationships, ensuring that knowledge of the external world is always yoked to the needs and motivations of the observer.

In chapters 10–13, which are admittedly more speculative, Edelman applies his theory to a host of difficult psychological problems, from the relation between language and thought to the thorny problems of unconsciousness and attention, and even to disorders of consciousness. Edelman might well have chosen to write a book on any one of these topics had he wished to argue for his theory in the kind of detail devoted to perception and memory here. However, his goal is merely to demonstrate that his selectionist approach can generate “in principle” models as workable, or better, than existing instructionist models. I was particularly impressed with chapter 13, “Diseases of Consciousness”, in which Edelman shows how a number of apparently diverse symptomatology can be understood in terms of local interruptions of re-entrant networks which yield the kind of peculiar dissociations of function and pathology so characteristic of neurological disease.

Despite its great strengths and suggestiveness, there are several important difficulties facing the theory. Edelman ties himself to accounting for consciousness in terms of evolutionary homologies. Thus, for example, no invertebrate can be said to have even the components of primary consciousness (such as global mappings) simply because invertebrates lack the requisite anatomy. But what about evolutionary analogies? Could not cephalopods, for example, have independently evolved some of the higher psychological functions? Here the parallel with immunobiology is again instructive: invertebrates cannot have ‘true’ immune functioning, but many have independently evolved systems which accomplish most of the true vertebrate immune functions. Can Edelman’s selectionist theory offer us equally valuable guidelines for assessing evolutionary convergence? Little is said to give guidance, and there are several worrisome oversimplified ‘adaptionist’ arguments for the evolution of consciousness.

The most serious weakness of the theory is the tendency to treat higher-order consciousness as separate from perception. Edelman here misses an important consequence of his own theory: if perceptual categories result in part from patterns of exploratory activity, then even the ‘higher’ senses of vision and hearing are proprioceptive as well as exteroceptive. Vision in adult humans is probably the most powerful proprioceptive modality. Edelman correctly wants to avoid confusing the proprioception of one’s body’s behaviour with awareness of one’s self and its action. But if the goal objects of vision are integrated with internal need and

motivation systems (as Edelman proposes in his account of the “cortical appendages”), then the higher-order self would almost surely emerge from conceptual reflection concerning the process of ambulatory perception. Thus, the connection between primary and secondary consciousness is probably more intimate than Edelman believes. To put it another way, his book is as much about perceiving one’s self acting in the world as it is about “the remembered present”.

This final book in Edelman’s trilogy is just what the doctor ordered for modern neuroscience: an attempt to explain important psychological facts without explaining them away, and without restricting the analysis to limited levels of structure of function — an attempt, in short, to offer a real, full-blown, falsifi-

able theory, not just another model or two. This hubris on Edelman’s part will entice many into attempting to prove him wrong. But without grand theories that can be proven wrong, sciences cannot advance. In his posthumous *Outlines of Psycho-Analysis*, Freud wrote that “The starting point of this investigation is provided by a fact without parallel, which defies all explanation or description — the fact of consciousness.” By rejecting both Freud’s early ‘gnostic-unit’ assumption and his later dualism, Edelman has begun the arduous task of explaining consciousness scientifically. □

Edward S. Reed is in the Drucker Brain Injury Center, Moss Rehabilitation Hospital, and the Department of Humanities and Communications, Drexel University, Philadelphia, Pennsylvania 19104, USA.

By the book

Stuart Orkin

Molecular Cloning: A Laboratory Manual, 2nd edn. Three volumes. By J. Sambrook, E. F. Fritsch and T. Maniatis. Cold Spring Harbor Laboratory Press: 1989. Spiral bound \$115.

FEW molecular biologists welcome publication of any of the many protocol books that promise to be the single source for their laboratory methods. For the most part, such manuals fall far short of this goal. So why the excitement surrounding the long-awaited second edition of the classic guide, *Molecular Cloning*, which first appeared in 1982? The original version immediately filled the need for an anthology of laboratory procedures pertinent to the emerging field of recombinant DNA. With the 545-page spiral-bound paperback in hand, virtually any experimentalist could make a stab at cloning and have a reasonable expectation of success. In the ensuing years, students, fellows, and faculty alike have found the format pleasing, the background descriptions of methods concise and informative, and the procedures generally easy to follow and execute. A generation of molecular biologists grew up with “Maniatis”, as this ‘bible’ was fondly dubbed.

Much has changed since *Molecular Cloning* first appeared on the scene. Recombinant DNA technology has greatly expanded and matured. Many new methods for the analysis and cloning of nucleic acids have been developed. Strategies for the expression of cloned genes, dissection of their controlling elements and characterization of interactions with regulatory proteins all post-date the first edition. Merely to remain contemporary, laboratories must now be familiar with methods traditionally under the purview

of protein chemists, cell biologists, bacterial geneticists and immunologists. The 1982 version of *Molecular Cloning* has thus long been out-of-date, particularly in those laboratories dedicated to molecular biology. At least two recent compendia, *Guide to Molecular Cloning Techniques* (edited by S.L. Berger and A.R. Kimmel) and *Current Protocols in Molecular Biology* (edited by F.M. Ausubel *et al.*), have superseded the first edition of Maniatis *et al.*

The new edition of *Molecular Cloning* is indeed an impressive achievement. Its growth in size alone bears witness to the explosion in molecular techniques. The original single volume has sired three, weighing more than 8 lb in total. This expansion reflects more comprehensive coverage of topics previously included in the first edition, as well as the addition of new chapters, such as those on oligonucleotide probes and mutagenesis, *in vitro* amplification by the polymerase chain reaction, expression of cloned genes in *Escherichia coli* and mammalian cells, and analysis of proteins expressed from cloned genes. Particularly noteworthy are the sections dealing with construction of complementary DNA and genomic libraries. Current tendencies to ‘clone by kit’ and ‘clone by phone’ aside, these sections provide thoughtful accounts of different approaches with a balanced view of their inherent strengths and weaknesses. As with the previous edition, the protocols are presented in a pleasing, easy-to-read format. Inclusion of the index in only the third volume is inconvenient, as the three volumes are unlikely to remain together in a busy laboratory environment.

Although quite wonderful, the second edition of *Molecular Cloning* is not perfect. That, of course, would be too much to ask. As expression cloning of cDNA and the cloning of large DNA fragments in yeast artificial chromosome vectors are particularly powerful recent methods, I

would have favoured inclusion of specific protocols, rather than mere mention of each. Similarly, procedures for the generation of subtracted cDNA libraries, for production of recombinant retroviruses and for amplification of expressible cDNA in mammalian cells are lacking. Given exciting developments in the study of regulatory proteins, I was surprised to find that no protocols for the analysis or purification of DNA-binding proteins were included. Although such techniques may fall beyond molecular cloning *per se*, they are as relevant to current research as the immunological methods described for the analysis of proteins expressed from cloned genes.

The rapid growth of recombinant DNA technology has created one thorny problem: it is no longer possible for a single protocol manual of manageable size to encompass the field. At present, the two best guides are *Current Protocols in Molecular Biology* and the second edition of *Molecular Cloning*. Either will stand up well to nearly every test in an active laboratory. Although presented in an unwieldy, loose-leaf format, *Current Protocols* deals effectively with rapid change in methodology by providing (at a cost) supplements, many of which concentrate on areas not covered in *Molecular Cloning*, such as growth and handling of yeast, *in situ* autoradiography, and protein-DNA methods. On the other hand, *Molecular Cloning* is of truly exceptional value. By virtue of its heritage and modest cost, it is likely to be more widely used by students at all stages of development.

Sambrook, Fritsch and Maniatis have done a great service to the molecular community, and deserve our gratitude. Although manuals of this kind cannot by themselves transform readers into innovative molecular biologists, they do serve to codify basic methods and simplify common procedures. Subsequent editions of *Molecular Cloning*, if forthcoming, will need to be encyclopaedic in scope to fill current gaps and account for future developments. Perhaps a computer-based format will be the final evolution of this literary genre. Otherwise, students of molecular biology should go into physical training right now merely to carry the protocol volumes of 1995 around the laboratory! □

Stuart Orkin is in the Division of Hematology/Oncology, Howard Hughes Medical Institute, Children's Hospital, Boston, Massachusetts 02115, USA.

■ For details of other available laboratory cloning manuals, see the review by Frank Gannon, John Neilan and Richard Powell in *Nature* 333, 309–310; 1988. One of the manuals discussed by Gannon *et al.* and mentioned above is now available in a revised and shortened form as *Short Protocols in Molecular Biology*, edited by F. M. Ausubel *et al.* Published by Wiley, price \$31.65, \$39.95.

From fertile soils

Valery N. Soyfer

The Monogerm of the Sugar Beet: Embryology, Genetics, Breeding. By S. I. Maletsky, Yu. N. Shavrukov, S. G. Veprev, E. I. Maletskaya, O. A. Kudryavtseva, A. V. Mglinets & M. A. Kostirya. *Novosibirsk, Nauka: 1988 (in Russian). Pp.168. 1 rouble 80 copeck.*

ALL Soviet science suffered from Lysenko and communist ideology, but Soviet plant geneticists suffered the most. Lysenko considered himself to be a specialist in the field of plant biology (it was not until the 1960s that he moved on to livestock breeding). He made every effort to exclude from the Soviet scientific community any scientist, breeder or agronomist who used the principles of genetics in his work. The terrible results of this policy were soon realized. The Soviet study of plant



Trofim Denisovich Lysenko — whose activities were among the factors leading to the breakdown of biological, agricultural and, to some extent, medical science in the Soviet Union.

genetics, which had until 1948 made outstanding achievements, was eliminated. When the taboo on genetics research was lifted after 1964, Soviet plant geneticists had no roots. This was particularly the case in sophisticated fields such as plant biotechnology, the genetic principles of development, self-incompatibility and cytoplasmic heredity. But now the first sprouts of contemporary genetics research has appeared in the shape of this book.

Stanislav Maletsky is a head of laboratory in the Institute of Cytology and Genetics of the Siberian branch of the USSR Academy of Sciences. His book, written with six colleagues, bridges the gap between lysenkoism and contemporary science. Beginning in the eighteenth century, and lasting until the 1960s, sugar-beet varieties used in agriculture were grown only from plants of so-called multigerm lines. The flowers of these plants occur in one cluster which, when sown, produce several seedlings. As a result, extra seedlings need to be thinned out by hand to obtain an optimal number of plants.

Plant geneticists in the 1930s understood that it would be very useful to select strains of sugar beet, every plant of which would develop only one seedling from one seed (so-called monogerm sugar beet). The existence of a monogerm in sugar beet not only has a practical value; it also poses a very interesting and complex genetic question.

The first serious breakthrough in this direction was made by Vyacheslav F. Savitsky, a Russian emigré to the United States, who found a monogerm mutant of sugar beet and in the 1940s and 1950s developed from it a new line for plant breeders. As well as this technological revolution in sugar-beet production, Savitsky also postulated a hypothesis concerning the role of *M-m* genes in determining the mono- and multigerm phenotypes of sugar plants. In 1964, K. Sedlmayr, a leading plant geneticist, wrote in a review:

V. F. Savitsky started his search of separately flowering mutants at the time when such a search seemed to be meaningless. But he was lucky and found several such mutants. They paved the foundation of the whole American program for the selection of monogerm varieties. . . . Only his fanatical endurance and genetical flair could lead to success: he was rewarded — he found the rare mutants: he propagated them, and he used them in plant breeding.

Thereafter, agriculture in the United States and Western Europe used only this one source of monogerm sugar beet. All attempts to find other mutant lines were unsuccessful. But because of the Iron Curtain, Soviet and Eastern European plant geneticists had to search for new mutants. After 10 years, Maletsky found two new mutant alleles of the sugar beet with separate female and male flowers. He studied in detail these new alleles of the *M-m* locus and bravely investigated self-incompatibility of sugar beets and the usefulness of heterosis (hybrid vigour) which is obtained when these self-incompatible lines are crossed. It took him 10 years of intensive work to obtain several inbred lines of sugar beet with monogerm character.

Maletsky and his co-authors consider the general problem of monogermity in their monograph, which gives a synthesis of the pertinent literature covering the past 50 years. The authors collect and discuss data on embryology, genetics and breeding of the monogerm sugar beet. But there is one important difference here from most Soviet scientific books: the authors do not overestimate Soviet successes or underestimate Western results. The book is thus a well-balanced description of study of this important problem.

The authors' attempted comparison between Soviet and Western data provides an opportunity to study how different

approaches were used to reach the same result. I believe that this is a unique book in that the Soviet authors openly explain how, after Lysenko forbade all genetics research in the Soviet Union in 1948, Soviet specialists drastically changed their methods of the study of monogermity. Instead of pure genetic work, they undertook a gigantic project to search for plants with flowers of predominantly male or female type. This search was organized to exclude genetic methods. Next, various lines with separately flowering plants were cross-pollinated and individual forms giving monogerm progeny were selected. The descriptions of these massive numbers of crossings and their results are of great interest for Western scientists. It is ironic to read that, paradoxically, despite the Soviet scientists' desire to turn away from genetics, their work resulted in a typical genetics investigation.

The central part of the book is devoted to Savitsky's hypothesis of the role of the

$M-m$ locus in the determination of the mono- and multigerm character of sugar beet. The most interesting part here is the detailed explanation of the screening procedures used by Soviet specialists during the search of the new representatives of $M-m$ alleles which enabled them to select and study in detail four new alleles of this locus in different varieties of sugar beet.

Western plant geneticists will be interested in this book because it represents a unique collection of Soviet data over a long period of time. Because the authors did not limit themselves to a discussion of their own work, but have tried to compare their data with the results of Western researchers, it is particularly informative. It would be of even more interest if the book were to be translated into English. □

Valery Soyfer is a Distinguished Visiting Professor in the Department of Molecular Genetics and the Ohio State Biotechnology Center, Ohio State University, Columbus, Ohio 43210-1292, USA.

An arrow's point

Robert M. Wald

The Physical Basis of the Direction of Time. By H. D. Zeh. Springer-Verlag: 1989. Pp.166. £6.50, DM18.

THE nature of time irreversibility in the laws of physics and in physical phenomena remains one of the most intriguing fundamental issues in physics. The presently known dynamical laws of physics are exactly CPT invariant, and the time-reversal asymmetry inferred from observed CP violation is extremely small. Nevertheless, most macroscopic physical phenomena we observe in our daily lives display a gross time asymmetry. Most prominent among these observed "arrows of time" is the "thermodynamics arrow" represented by the law of increase of entropy. In addition, there is the obvious fact that our perceptions of the past and future are totally different — we cannot 'remember' the future. Indeed, it takes considerable training in physics to learn to treat the 'past' and 'future' on a comparable footing.

The (nearly exact) time symmetry of the laws of physics is not necessarily in any conflict with the "arrows of time" observed in physical phenomena: time-symmetrical dynamical equations admit time-asymmetrical solutions. In particular, the thermodynamic arrow of time can be attributed to 'special' initial conditions for our Universe. Nevertheless, most physicists expect (or at least hope) that a more fundamental origin of the observed arrows of time eventually will be uncovered, and that some deeper relationships between them will be found.

It seems clear that some arrows of time — such as the retardation of radiation — are closely related to the thermodynamic arrow. It is less clear that there is a relationship between the thermodynamic arrow and the arrow of time in quantum measurement theory associated with reduction of the state vector. (Indeed, it is debatable whether there really is an arrow of time in quantum measurement theory, although Penrose has recently provided a strong argument for the presence of such an arrow.) It is far less clear that the arrow of time associated with the expansion of the Universe bears any relationship to the other arrows. The ideas and speculations of about 10 years ago on the nature of the various arrows of time and the relationships between them were discussed and summarized by Paul Davies in his book *The Physics of Time Asymmetry*. Since then, there have been several noteworthy developments, mainly arising from research in general relativity and quantum gravity. In particular, Penrose has proposed that the vanishing of the Weyl tensor at initial singularities could account for the thermodynamic arrow. In quantum cosmology, proposals have been advanced for boundary conditions which determine the wavefunction of the Universe; the relationship between the expansion of the Universe and the other arrows of time can then be addressed, at least in the context of the highly simplified and idealized models for which the proposals can be given concrete form.

Zeh provides a concise, technically sophisticated and up-to-date discussion of research on the arrows of time in physics. He displays a wide familiarity with the relevant modern literature in the rather diverse fields of study related to the arrows of time. Indeed, one of the most

valuable aspects of his book is the bibliography of the original research papers. However, much of the book reads as a terse commentary on these papers. The author frequently attempts to summarize sophisticated and controversial ideas in a sentence or two, with the result that the reader will have little chance of following the discussion unless he already is thoroughly familiar with these ideas. In addition, although I did not encounter any serious technical errors, there are several slight misstatements and other minor flaws, particularly in the last two chapters on space-time structure and quantum cosmology. (An example of such a flaw is the failure to draw a distinction between Penrose's (time symmetrical) strong cosmic censor hypothesis — which postulates that space-times must be globally hyperbolic, thereby allowing one to distinguish between 'past' and 'future' singularities — and Penrose's (time asymmetrical) Weyl tensor hypothesis — which postulates that the Weyl tensor must vanish near past singularities.) These flaws contribute to the difficulty in following some of the arguments.

Zeh's main thesis seems to be that the expansion of the Universe provides a "master arrow of time" from which all others follow. (However, this conclusion is stated more clearly and explicitly on the jacket than anywhere in the book.) The basic idea behind this suggestion has to do with the fact that in minisuperspace models, the radius of the Universe formally plays the role of 'time' in the Wheeler-DeWitt equation satisfied by the wavefunction of the Universe. If one imposes boundary conditions that select real solutions of the Wheeler-DeWitt equation, then — at least in some formal sense — there should be a symmetry between expanding and contracting phases in the evolution of the Universe and, hence, all other arrows of time should correlate with the expansion of the Universe.

A similar conclusion was drawn by Hawking several years ago, although he retracted this conclusion after it was recognized that most classical trajectories associated with the wavefunction of the Universe probably would not display a symmetry between the expanding and contracting phases. My own view is that there are sufficiently many unresolved interpretive issues concerning the wavefunction of the Universe that it would be premature to draw any physical conclusions at the present time. Despite these criticisms, however, Zeh's book provides a very useful guide to modern research on the nature of the arrows of time in physics. □

Robert M. Wald is in the Enrico Fermi Institute, University of Chicago, 5630 Ellis Avenue, Chicago, Illinois 60637, USA.

Abrupt climate fluctuations in the tropics: the influence of Atlantic Ocean circulation

F. Alayne Street-Perrott & R. Alan Perrott

Several prolonged droughts in the Sahel and tropical Mexico during the past 14,000 years were coincident with large injections of fresh water into the northern North Atlantic Ocean. The link between these phenomena lies in the thermohaline circulation of the oceans: input of fresh water decreases salinity leading to reduced North Atlantic Deep Water formation and anomalies of sea surface temperature of the kind associated with decreased rainfall in the northern tropics. Ice-sheet disintegration, the most important source of fresh-water input to the oceans, should therefore be considered explicitly in models of past and future climate.

ONE of the greatest climatological puzzles of the twentieth century has been the severe drought in the Sahel over the past 20 years^{1,2}. After almost five decades during which annual rainfall totals in sub-Saharan Africa hovered near or slightly above the long-term mean, drought set in 1968 and continued with little respite until 1988. Recent studies have established a statistically significant relationship between below-normal rainfall in the Sahel and a persistent, asymmetrical pattern of global sea surface temperatures (SSTs), characterized by colder-than-average conditions north of the Equator and warmer-than-average waters to the south³⁻⁵. Newell and Hsiung⁴ suggested that this pattern is at least partly generated by a reduction in the northward transport of heat by the Atlantic Ocean.

During the past 14,000 years, lake levels and other sensitive indicators of tropical rainfall have also recorded dramatic fluctuations, superimposed on the slow trends attributable to orbital forcing. These abrupt changes were most pronounced in the areas affected by recent Sahel droughts, but endured for much longer—centuries to millennia.

Here, we draw on a large body of recent oceanographic and palaeoceanographic literature to argue that several severe historical and prehistoric droughts were linked to episodes of decreased salinity in the northern North Atlantic. Freshening tends to stabilize the water column, thereby inhibiting deep convection and reducing the compensating northward transport of heat across the Equator by the thermohaline circulation. There is considerable controversy about the role of glacial meltwater in the major low-salinity events recorded during the late glacial and early post-glacial. Similar feedback effects, triggered by increased precipitation at high latitudes or by rapid

meltback of the Greenland Ice Sheet, could give rise to 'disagreeable surprises' during the transient response of the climate system to a future greenhouse warming^{6,7}.

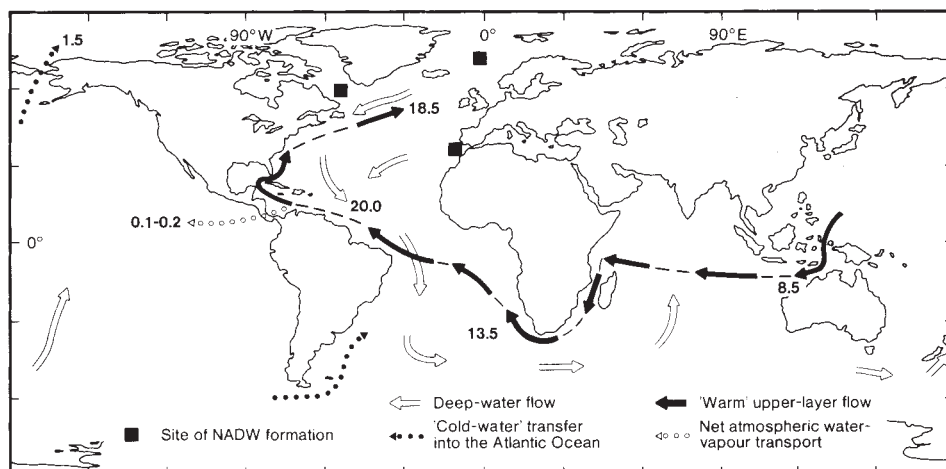
Sahel drought and SST anomalies

Significant progress in the understanding of droughts in sub-Saharan Africa has been made through the statistical analysis of large sets of SST observations compiled from ship reports^{3,8-11}. Parker *et al.*⁵ carried out an eigenvector analysis of a carefully standardized, global database of seasonal SSTs for the period 1901–80. The dominant spatial pattern identified, EOF1, represented the global trend in mean SST, whereas the second, EOF2, was dominated by SST variations in the tropical East Pacific. EOF3 was highly correlated with the Sahelian rainfall index of Nicholson¹; below-normal rainfall was associated with cold anomalies in the North Atlantic and North Pacific, and warm anomalies in the South Atlantic, Indian Ocean and southwest Pacific.

The heating anomalies resulting from these large-scale SST departure fields were surprisingly large. The mean modulus of the changes in July–September between a composite of five Sahel drought years and five wet years was 15 W m^{-2} averaged over the World Ocean⁵. This figure can be compared with the estimated changes in radiative heating at the Earth's surface due to a doubling of carbon dioxide (4 W m^{-2} ; ref. 12), and the increased summer insolation over the Northern Hemisphere 9,000 years ago (20 W m^{-2} ; ref. 13).

The north–south oscillation in SSTs represented by EOF3 seems to modulate the intensity of the moisture convergence into the Intertropical Convergence Zone and to a limited extent,

FIG. 1 Global thermohaline circulation cell associated with NADW production (modified after ref. 21). Solid arrows mark the inferred 'warm-water route' for return flow of upper-layer water to the northern Atlantic. Suggested volume fluxes (Sv)²¹ are based on uniform upwelling of NADW with a production rate of 20 Sv. The alternative 'cold-water route' through the Drake Passage is assumed to be minor. The atmospheric water-vapour flux across the Isthmus of Panama is estimated from data in refs 19, 20.



the latitude of the region of strongest convective activity within it^{3,14}. The particular sensitivity of sub-Saharan rainfall to SST may, however, also reflect the strong positive feedbacks between the monsoon rains and local surface boundary conditions such as soil moisture¹⁵.

Although the unusual persistence of the recent Sahel drought is often attributed to human-induced desertification¹⁶, EOF3 has exhibited strong temporal trends. A marked north(cold)–south(warm) contrast in SST characterized the entire period 1968–86 (ref. 5). Such a long-lasting anomaly pattern of world-wide extent is explained most readily by a change in heat transport by the general oceanic circulation. The change in the energy content of the Atlantic alone between the wet period 1949–64 and the dry period 1965–79 implies a reduction of 1.2×10^{14} W ($\sim 20\%$) in the net northward flux of heat crossing the Equator⁴.

The thermohaline 'conveyor belt'

At present, the North Atlantic and North Pacific are strikingly asymmetrical. Between 45° and 60° N, the North Atlantic has an annual mean SST of 10° C and a salinity of 34.9‰, whereas the North Pacific averages 6.7° C and 32.8‰, respectively¹⁷. The higher salinity of the North Atlantic reflects the net excess of evaporation over precipitation and runoff (~ 0.3 Sv between 0 and 80° N; ref. 18) (1 Sv $\equiv$ sverdrup $= 10^6$ m³ s⁻¹). Most of the excess evaporation occurs in the subtropical gyre; based on ref. 19, the net fresh-water loss between 10° and 25° N amounts to ~ 0.3 Sv, much of which is exported to the Pacific by the trade winds across the Isthmus of Panama (Fig. 1)^{17,20}.

Winter cooling in the Norwegian and Labrador Seas causes deep convective overturning, which generates North Atlantic Deep Water (NADW). The cold, saline NADW from these two areas is joined by the warmer but saltier outflow from the Mediterranean and flows southwards down the western Atlantic, eventually circulating around the World Ocean (Fig. 1). No deep water forms today in the North Pacific, where the evaporation rate is lower as a result of upwelling¹⁷.

The total production rate of NADW from all three sources on Fig. 1 is estimated²¹ as 15–20 Sv, leading to a heat gain by the atmosphere of the order of $5\text{--}7 \times 10^{14}$ W, or $\sim 20\text{--}30\%$ of the solar energy received by the Atlantic Ocean north of 35° N (ref. 18). According to Gordon's model²¹, the oceanic heat loss is compensated by heat transport in the global thermohaline circulation cell shown in Fig. 1. Warm, relatively fresh North Pacific thermocline water is driven through the Indonesian straits into the Indian Ocean, mainly during the northern summer monsoon season²², and then transported to the tip of South Africa in the south equatorial and Agulhas currents. During this transit, it gains volume by upwelling, and salt by excess evaporation²¹. Warm eddies from the Agulhas current spin off into the South Atlantic, transporting up to 5×10^{14} W of heat²³. This energy is carried northwards by the Benguela current, passing across to the western North Atlantic in the complex equatorial circulation. Estimates of the mean annual heat flux across the Equator^{24–26} range from 5 to 14×10^{14} W.

The present global thermohaline circulation seems to be self-sustaining²⁷. The transport of water vapour by the atmosphere from the Atlantic to the Pacific is compensated by the large-scale circulation of salt-laden NADW, and the deep outflow from the North Atlantic is balanced by a warm surface inflow which maintains the anomalously high evaporation rate. But according to recent experiments²⁸ using a coupled atmosphere–ocean general circulation model (GCM) with realistic geography, there may be two quasi-stable equilibria, one with and one without a strong thermohaline circulation in the Atlantic (typified by experiments I and II of ref. 28, respectively). In the latter, NADW production was suppressed by a strong halocline. SSTs were colder than in experiment I in the North Atlantic and to a lesser extent in the North Pacific, but warmer in the three southern oceans; cross-equatorial heat transport was reduced

by $\sim 4 \times 10^{14}$ W; and the tropical rainbelt shifted southwards, giving rise to decreased precipitation over West Africa and the North American tropics. The ocean circulation could be switched to this alternative state by a significant decrease in the salinity of the North Atlantic^{20,29,30}. Such a freshening might be induced by a substantial reduction in the evaporation excess in the subtropics²⁰, or by a large input of fresh water in the form of melt water, river runoff or outflow from the Arctic Ocean²⁹.

A significant freshening of the northern Atlantic, known as the 'Great Salinity Anomaly', was observed³¹ in 1968–82. The deep water in both the Norwegian and Labrador Seas became significantly cooler, fresher and less oxygenated³², in keeping with a decrease in surface salinity of 0.1–0.2‰ (refs 31, 32). The cause of these changes may have been anomalous northerly flow over the Greenland Sea³² or a strong upward trend in precipitation observed over land areas between 35° and 70° N during the past 30–40 years³³. Bradley *et al.*³³ speculated that the prolonged recent sub-Saharan drought and the increase in precipitation at mid to high latitudes were both signs of the transient response of the climate system to rising levels of carbon dioxide. An outbreak of low-salinity water from the Arctic around 1908–14

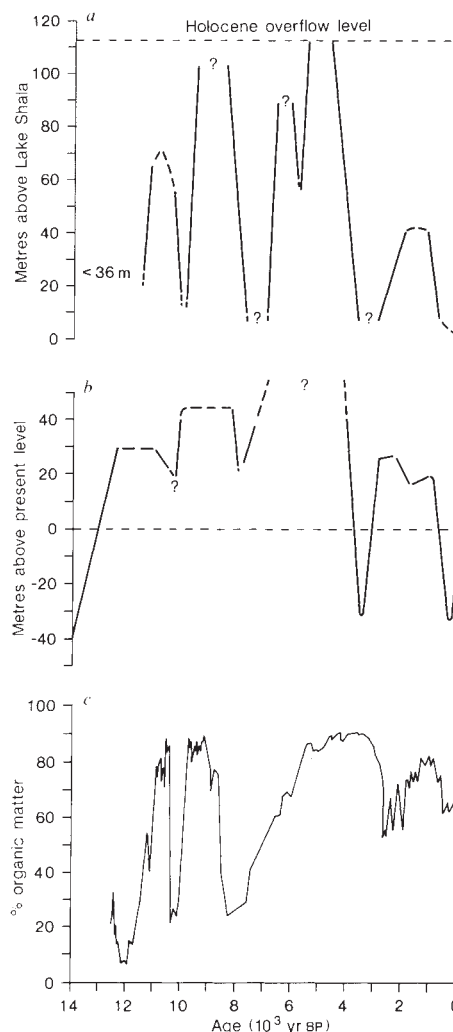


FIG. 2 a, Fluctuations in lake level, Ziway-Shala Basin (Ethiopia), 14,000–0 yr BP ($7^\circ 45'$ N, $38^\circ 40'$ E, 1,558 m above sea level) (ref. 39); b, Fluctuations in lake level, Lake Bosomtwi (Ghana) ($6^\circ 30'$ N, $1^\circ 25'$ W, 100 m above sea level) (ref. 46); c, Variations in organic content, Mt Satima mire (Kenya) ($0^\circ 18'$ S, $36^\circ 35'$ E, 3,670 m above sea level) (R.A.P., unpublished data).

(ref. 31) was, however, also coincident with severe drought in sub-Saharan Africa¹.

The upper branch of the thermohaline circulation cell shown in Fig. 1 could be affected by changes well outside the confines of the North Atlantic. For example, the flow through the Indonesian Archipelago may have decreased in glacial times in response to lower sea levels and to the decreased strength of the East Asian summer monsoon³⁴. The Agulhas current transport around the Cape of Good Hope is sensitive to the latitude of the subtropical convergence^{21,35}. Whatever the cause of the perturbation, however, the rate of deep-water formation could respond on timescales much shorter than the turnover time of the deep ocean. The recent freshening of the North Atlantic had an impact within a decade^{31,32}. In GCM experiments^{7,30}, significant changes in NADW production have been simulated on a timescale as short as $10\text{--}10^2$ yr.

We now turn to the record of changes in tropical precipitation during the past 14,000 years.

Abrupt tropical climate events

Fluctuations in the extent and depth of closed lakes provide some of the most widely available and closely dated evidence for climate changes during late Pleistocene and Holocene times. The response of any individual lake depends on its hydrology³⁶. A closed, amplifier lake (the type most sensitive to climate change) receives the bulk of its inputs in the form of surface runoff and loses water only by evaporation, provided that losses to ground water are negligible. The rate of response of its water surface area, A_1 , to an abrupt change in climate is governed by its characteristic response time (e-folding time) τ_e , given by $\tau_e = A_1 / [(dA_1/dL)(E_1 - P_1)]$, where D is lake depth and E_1 and P_1 are annual evaporation and precipitation over the water surface, respectively³⁷. In the dry and seasonal tropics, τ_e tends to be short because of large values of $(E_1 - P_1)$; for lakes there with areas of $10^2\text{--}10^3$ km², τ_e is generally of the order of $10\text{--}10^2$ yr.

Closed amplifier lakes in the tropics are therefore excellent indicators of variations in water balance on a timescale of $10^2\text{--}10^3$ yr (ref. 38). Figure 2a shows the record for Lake Ziway-Shala (Ethiopia) for the past 14,000 years³⁹. This curve, which is based on a large number of dated shorelines and sedimentary sections, is controlled by 40 ¹⁴C dates. Low levels prevailed before 12,000 radiocarbon years before present (yr BP) and in the intervals $>10,400\text{--}9,800$, $7,800\text{--}7,200$ and $\sim 4,500\text{--}2,500$ yr BP. There was also a minor recession centred on 5,900 yr BP. Water levels have remained low during the past 1–2 millennia. This chronological pattern of fluctuations is typical of amplifier lakes in eastern Africa. Similar sequences have been obtained from Lakes Abhé (11°N)^{40,41}, Turkana (5°N)^{42,43}, Nakuru-Elmenteita (0°S)⁴⁴, and Rukwa (8°S)⁴⁵, although individual fluctuations are not equally well resolved in all sequences. For example, the recessions dated 11,000–10,000 and 8,000–7,000 yr BP seem to become less marked to the south.

Figure 2b shows the water level curve for Lake Bosumtwi (Ghana), which is representative of West Africa south of the Sahara. The sequence is based on a detailed stratigraphical and mineralogical analysis of cores and exposed sections, supported by 42 ¹⁴C dates⁴⁶. It shows lake level minima before 13,000 yr BP, and during the intervals 11,200–10,100, 8,200–7,100, 4,200–3,000 and 1,000–0 yr BP. Low stands occurred at approximately the same times in Lake Chad (13°N)⁴⁷.

Further evidence for abrupt climate events is provided by high-resolution stratigraphies from aeolian, fluvial and swamp deposits^{48–51}, although their importance has not always been recognized. Figure 2c shows the variations in the organic content of a high-altitude mire on Mt Satima (Kenya) (R. A. Perrott, unpublished data). Low values record periods of slow peat growth when the bog surface was dry. Based on a total of 21 ¹⁴C dates, peat development began around 12,000 yr BP, but

was interrupted by episodes of desiccation during the intervals 10,500–9,900, 8,500–6,500 and $\sim 2,800$ yr BP.

Rapid lake level fluctuations also occurred during the past 14,000 years in the North American tropics. Figure 3 shows the relative water level variations in Lake Chiconahuapan (upper Río Lerma Valley, Mexico) (ref. 52 and S. E. Metcalfe, unpublished data), and in the Chalco embayment of the Basin of Mexico, just to the east⁵³. Both basins are fed partly by springs and partly by surface runoff. The lake level index for Lake Chiconahuapan is based on the relative proportions of planktonic, facultative-planktonic, epiphytic and aerophilous diatoms in two pit sections⁵⁴. The ¹⁴C chronology comprises 10 dates on organic matter. In contrast, the Chalco sequence was derived from sedimentological and palaeoecological investigations of the Tlapacoya archaeological site⁵³. This curve is fixed by nine ¹⁴C dates. Two regional tephra^{55,56} are also used in Fig. 3 to correlate the two sequences.

Although the two Mexican curves exhibit differences that may reflect their relatively coarse dating framework as well as lags introduced by groundwater flow, they have important common features. A high stand of Lake Chalco, following several millennia of low water levels, began after 12,900 yr BP and peaked just before the eruption of the Upper Toluca Pumice, $\sim 11,600$ yr BP. It probably correlates with undated organic sediments that underlie the same pumice layer in Lake Chiconahuapan. In contrast to the African sequences (Fig. 2), however, there is no evidence for a high stand between 10,000 and 8,000 yr BP, although the two Mexican curves are poorly dated in this interval. Maxima followed in both basins in the intervals 7,000–5,500 yr BP and 3,300–900 yr BP, separated by a phase of severe desiccation.

All the sites discussed so far lie in an area in which interannual variations in rainfall during the twentieth century have shown a strong positive correlation with those in the Sahel^{57,58}. Variations in SST and atmospheric circulation over the tropical Atlantic seem to have been the linking factor¹⁰. Although the general occurrence of high lake levels in the northern tropics

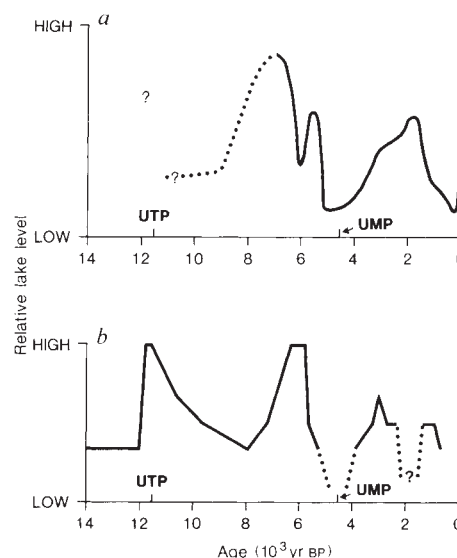


FIG. 3 a, Relative fluctuations in the level of Lake Chiconahuapan (Mexico) 14,000–0 yr BP ($19^\circ 08'\text{N}$, $99^\circ 31'\text{W}$, 2,575 m above sea level) (ref. 52 and S. E. Metcalfe, unpublished data); b, Relative fluctuations in the level of Lake Chalco (Mexico) ($19^\circ 24'\text{N}$, $99^\circ 11'\text{W}$, 2,250 m above sea level) (ref. 53). Key volcanic marker beds: UTP, the Upper Toluca Pumice, firmly dated at $11,580 \pm 70$ yr BP ($n=4$) in the upper Río Lerma Valley and $11,560 \pm 70$ yr BP ($n=3$) at Tlapacoya (ref. 55); UMP, the Yellow Ash/Upper Marker Pumice, dated at $4,570 \pm 60$ yr BP in Lake Chiconahuapan, Pit 1 (ref. 52) and between 4,880 and 4,250 yr BP at Tlapacoya (refs 53, 56).

between ~15,000 and 6,000 yr BP can be explained by the maximum of summer insolation caused by the Earth's orbital configuration^{13,59}, the rapid climate shifts described above cannot. In the discussion that follows we focus on the changes in ocean circulation associated with the most marked of these abrupt events: (1) the sudden rise in lake levels around 13,000–12,000 yr BP after 3–4 millennia of low or intermediate water levels; (2) the rapid fall around 11,500–11,000 yr BP, which was followed by an equally fast recovery at ~10,000 yr BP in northern intertropical Africa but not in Mexico; and (3) the episodes of low lake level during the intervals 8,000–7,000, 5,000–3,000 and 1,000–0 yr BP.

Palaeoceanographic evidence

How did the thermohaline circulation change during the glacial-interglacial transition? A pattern of north-south contrast in tropical Atlantic SSTs during the past 20,000 yr has been identified by Mix *et al.*⁶⁰, who carried out an eigenvector analysis of SST estimates based on planktonic foraminiferal assemblages in deep-sea cores. EOF2, which accounted for 12% of the total variance, represented relatively cool conditions in the North Atlantic and warm conditions in the South Atlantic. It exhibited two peaks at roughly 14,000–13,000 and 11,500–10,500 yr BP. They attributed this mode to a reduction in the cross-equatorial heat flux in response to the suppression of thermohaline overturning by glacial meltwater⁶⁰.

A reduction in deep-water formation during the last glacial is suggested by the assemblages⁶¹, stable-isotope ratios^{62–65} and trace-element chemistry⁶⁶ of bottom-dwelling foraminifera in Atlantic cores. The most compelling evidence consists of variations in the $^{13}\text{C}/^{12}\text{C}$ and Cd/Ca ratios. Although $^{13}\text{C}/^{12}\text{C}$ (ref. 67) and to a much lesser extent, Cd/Ca (ref. 68) may have been affected by glacial-interglacial changes in the total inventories of ^{12}C and Cd in the ocean, a large fraction of the variation in benthic foraminifera is attributable to changes in deep ocean circulation. Today, 'older' water masses have lower ^{13}C values and higher nutrient contents because of oxidation of ^{13}C -depleted organic matter in the deep water⁶⁹. Because cadmium is highly correlated with phosphorus in modern deep-water samples⁶⁸, the Cd/Ca ratio in benthic foraminifera from Atlantic cores also provides an index of the supply of 'young', oxygen-

ated, nutrient-depleted NADW. Using a simple model of deep-water composition, Boyle and Keigwin⁶⁶ calculated that the intensity of this source relative to nutrient-enriched southern waters was halved during the last glacial maximum.

A record of much greater temporal resolution is now available from the western Atlantic⁷⁰. Core EN120 GGC1 from the Bermuda Rise (4,450 m depth) exhibits a threefold variation in Cd/Ca ratios between 14,500 and 0 yr BP (Fig. 4). At the beginning of the record, Cd/Ca was high, indicating greatly reduced NADW production. Deep-water formation resumed in the intervals 14,300–13,900, 12,800–12,200 and 12,000–11,000 yr BP, only to drop back abruptly towards glacial levels between 11,000 and 9,600 yr BP. Since then Cd/Ca ratios have remained low, although there were fluctuations that have not yet been confirmed by analyses on other cores with high sedimentation rates. This sequence⁷⁰ demonstrates that the deep ocean has undergone dramatic changes in circulation on a timescale of ≤ 500 years.

The Bermuda Rise sequence has recently been challenged by Jansen and Veum⁷¹, based on the benthic ^{13}C record of core V23-81, taken from the southern Rockall Trough just west of Ireland, which shows relatively high ^{13}C values during the Younger Dryas event (~11,000–10,000 yr BP). These are interpreted as indicating the resumption of deep convection following an episode of reduced ventilation centred on 13,700 yr BP. But the location of this site is far from ideal: it was probably prone to local melt-water effects, judging from the widespread distribution of distal glaciomarine sediments in the northern Rockall Trough⁷²; it is also relatively shallow (2,393 m), and is located well to the east of the main deep outflow from the Norwegian Sea through the Denmark Strait. Jansen and Veum concede that the record from V23-81 is compatible with a shallower and less extensive NADW mass during the Younger Dryas.

Discussion

The episodes of reduced NADW formation shown in Fig. 4 were linked to important events in the history of the northern ice sheets^{27,73}. From 18,000 until 14,000 yr BP their margins did not change greatly⁷⁴, apart from the disintegration of the marine-based Barents Sea ice sheet at ~15,000 yr BP (ref. 75). The northern Atlantic remained covered by pack ice⁷⁶. Dry

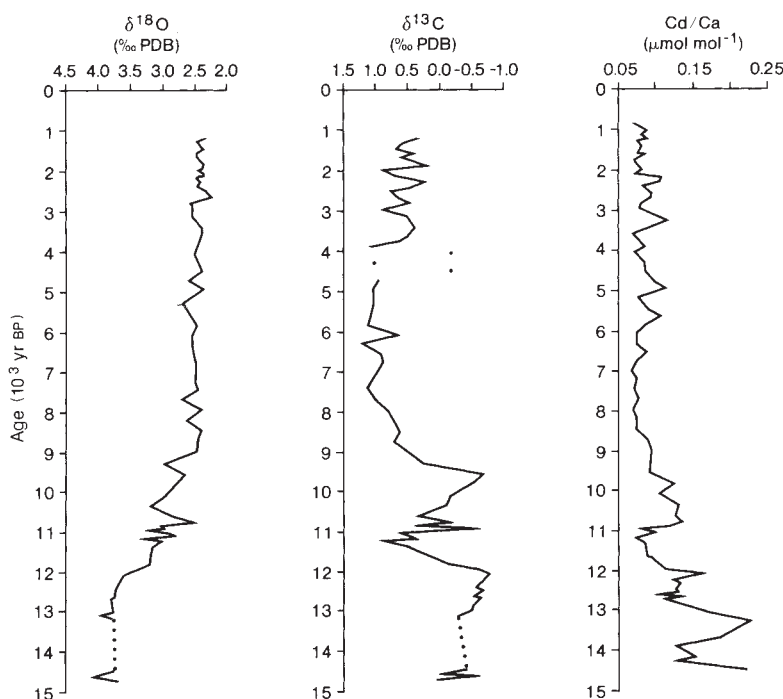


FIG. 4 Variations in benthic $\delta^{18}\text{O}$, $\delta^{13}\text{C}$ and Cd/Ca in ocean core EN120 GGC1⁷² plotted against estimated age. $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ are standardized departures of $^{18}\text{O}/^{16}\text{O}$ and $^{13}\text{C}/^{12}\text{C}$ from the PDB analytical standard in per mil (‰). The timescale is based on correlations with the nearby well dated core GPC5⁸¹.

conditions prevailed in the northern tropics. Rapid melting began around 14,000 yr BP, but initially the bulk of the melt water from the Laurentide ice sheet drained into the Gulf of Mexico through the Mississippi. The 'melt-water spike' in ^{18}O curves from the Gulf of Mexico, which is dated at ~13,600–11,100 yr BP (refs 27, 73), exhibits two peaks centred on 13,500 (smaller) and 11,800 yr BP (larger). During this interval, deep convection twice started up in the North Atlantic, halting briefly at the time of greatest outflow from the Gulf (Fig. 4). Lake levels rose sharply in northern intertropical Africa and Mexico.

As the Laurentide ice sheet retreated, melt water was increasingly channelled into the Gulf of St Lawrence, culminating in the eastward diversion of the outflow from the huge Glacial Lake Agassiz at ~10,800 yr BP (ref. 77). In the northern tropics, lakes fell back rapidly to glacial levels.

Broecker *et al.*²⁷ proposed that the sudden injection of large volumes of melt water into the northern Atlantic, just south of the Labrador Sea, was responsible for shutting off deep-water production during the Younger Dryas event. Previously, the sharp climate cooling experienced during this interval in the North Atlantic, northwest Europe and easternmost North America was attributed to the direct impact on the atmosphere of a catastrophic release of melt products, leading to cooling of the ocean and a re-advance of the pack ice^{76,78}. Now, two powerful reinforcing feedbacks have been proposed: a decrease in NADW production and a related drop in atmospheric carbon dioxide content¹⁸.

Minimum and maximum limits on the rate of melt-water injection into the northern Atlantic during the Younger Dryas are set by estimates of the baseflow from Lake Agassiz (0.03 Sv)⁷⁷ and the rise in eustatic sea level (0.09 Sv)⁷⁹. These amount to ~10% and 30%, respectively, of the evaporation excess that drives the modern thermohaline circulation. The eastward outflow from Lake Agassiz was terminated by an ice advance at 10,000 yr BP that diverted the melt water back into the Mississippi drainage for ~500 years⁷⁷. The North Atlantic became ice-free within 20 years⁸⁰.

Production of NADW resumed after 9,600 yr BP, reaching its maximum in the early Holocene^{70,71} (Fig. 4), when the summer monsoons were strongest in Asia and northern Africa in response to a 7% increase in insolation during the Northern Hemisphere summer^{34,59}. The Agulhas current flow may also have been greater than today³⁵. These changes are consistent with a slightly more intense thermohaline circulation. Although the absence of high lake levels between 10,000 and 8,000 yr BP in the North American tropics is puzzling, it is possible that the region of strongest convective activity was displaced northwards into southwest North America, as it is today at the height of the summer, in response to the larger heating anomaly over the North American land mass^{13,81}.

We attribute the arid episode recorded in Africa between 8,000 and 7,000 yr BP to the rapid evacuation of ice when the sea broke into Hudson Bay around 8,000 yr BP (ref. 82). Although the size of the Laurentide ice sheet just before this event is controversial⁸³, Flohn and Nicholson⁸⁴ estimated from the associated sea-level rise⁸⁵ that the release of melt water averaged 0.5 Sv over 200 years. This is a large enough perturbation to have completely suppressed deep-water formation, even though it would have been too short to be registered in deep-sea sediments.

Despite the coincidence of the two major melt-water pulses into the northern North Atlantic (10,800–10,000 and ~8,000 yr BP) with dry events in the tropics, several vexing questions remain to be answered. First, although recent GCM experiments^{86,87} show that a cold melt-water 'lid' in the Gulf of Mexico would result in lower temperatures and higher pressure over subtropical land areas, both periods of strong melt-water influx into the Gulf (around 13,600–11,100 and 10,000–9,500 yr BP) corresponded to significant increases in rainfall in northern intertropical Africa. It can be argued, however, that a

stable fresh-water layer in low latitudes would reduce the downward penetration of solar heating in summer and could actually result in increased SSTs⁸⁸, especially at a time of enhanced summer insolation^{13,59}.

Second, a recent study of submerged coral terraces off Barbados⁷⁹ indicates that the rate of eustatic sea-level rise peaked at 12,000 and 9,500 yr BP, with a significant decrease during the intervening Younger Dryas. Hence, to explain the reduction in NADW formation between 11,000 and 9,600 yr BP implied by the Cd/Ca data (Fig. 4), it is necessary to assume either that melt water injected directly into the northern North Atlantic was more effective in suppressing NADW formation than water entering the Gulf of Mexico, which could be trapped in the subtropical gyre, or that the two episodes of rapid sea-level rise⁷⁹ included large contributions from the breakup of more remote ice-sheet margins (such as Antarctica), which seems unlikely. It is also not clear why the reopening of the eastward outlets from Lake Agassiz around 9,500 yr BP (ref. 77) is not recorded in the Cd/Ca curve (Fig. 4). The chronological framework for EN120 GGC1 may not, however, be fine enough to distinguish this event from the Younger Dryas.

Third, the abrupt climate shifts experienced in the tropics since 6,000 yr BP cannot easily be attributed to variations in the ice sheets because there has been little glacio-eustatic sea-level change since that time⁷⁹. Although an external cause such as solar variations or volcanism cannot be ruled out, these events may have been due entirely to the kind of autovariation in the coupled atmosphere-ocean system exemplified by the 'Great Salinity Anomaly' of 1968–82. Whatever the cause(s), they had a large impact on the prehistoric cultures of tropical Africa⁸⁹ and possibly also of Mesoamerica⁹⁰.

Several new initiatives are required to address the important questions now emerging. Information from glacial geology, sea-level curves and isotope palaeosalinity records must be combined to derive a more quantitative and reliable picture of the temporal and spatial variations in glacial melt-water discharge. A closer network of high-sedimentation-rate deep-sea cores dated by accelerator mass spectrometry will be crucial⁷¹ to the understanding of changes in deep-water circulation in the Atlantic, especially in the critical depth range of 2,500–4,000 m. Greater emphasis needs to be given to numerical simulation of the impact of changes in buoyancy on coupled atmosphere-ocean models. Efforts to probe the climate changes of the past 6,000 years are also urgently required if the nonlinear behaviour of the climate system and, in particular, the possible effect of rising levels of carbon dioxide are to be properly understood.

Alayne Street-Perrott and Alan Perrott are at the School of Geography, Oxford University, Mansfield Road, Oxford OX1 3TB, UK.

- Nicholson, S. E. *J. Clim. appl. Met.* **24**, 1388–1391 (1985).
- Lamb, P. J., Peppler, R. A. & Hastenrath, S. *Nature* **322**, 238–240 (1986).
- Folland, C. K., Palmer, T. N. & Parker, D. E. *Nature* **320**, 602–607 (1986).
- Newell, R. E. & Hsiung, J. in *Abrupt Climatic Change* (eds Berger, W. H. & Labeyrie, L. D.) 67–87 (Reidel, Dordrecht, 1987).
- Parker, D. E., Folland, C. K. & Ward, M. N. in *Recent Climatic Change* (ed. Gregory, S.) 166–178 (Belhaven, London, 1988).
- Broecker, W. S. *Nature* **328**, 123–126 (1987).
- Stouffer, R. J., Manabe, S. & Bryan, K. *Nature* **342**, 660–662 (1989).
- Lamb, P. J. *Tellus* **30**, 240–251 (1978).
- Hsiung, J. & Newell, R. E. *J. phys. Oceanogr.* **13**, 1957–1967 (1983).
- Hastenrath, S. *Mon. Weath. Rev.* **112**, 1097–1107 (1984).
- Lough, J. M. *Mon. Weath. Rev.* **114**, 561–570 (1986).
- Ramanathan, V. *J. Atmos. Sci.* **38**, 918–930 (1981).
- Kutzbach, J. E. & Guetter, P. J. *J. Atmos. Sci.* **43**, 1726–1759 (1986).
- Palmer, T. N. *Nature* **322**, 251–253 (1986).
- Mintz, Y. in *The Global Climate* (ed. Houghton, J. T.) 79–105 (Cambridge University Press, 1984).
- Rasmussen, E. M. in *Drought in Africa: Denying Famine a Future* (ed. Glantz, M. H.) 3–22 (Cambridge University Press, 1987).
- Warren, B. A. *J. mar. Res.* **41**, 327–347 (1983).
- Broecker, W. S., Peteet, D. & Rind, D. *Nature* **315**, 21–26 (1985).
- Baumgartner, A. & Reichel, E. *The World Water Balance* (Elsevier, Amsterdam, 1975).
- Weyl, P. K. *Met. Monogr.* **8**, 37–62 (1968).
- Gordon, A. L. *J. geophys. Res.* **91**, 5037–5046 (1986).
- Murray, S. P. & Arief, D. *Nature* **333**, 444–447 (1988).
- Gordon, A. L. *Science* **227**, 1030–1033 (1985).
- Lamb, P. J. & Burkner, A. F. *J. phys. Oceanogr.* **12**, 1388–1410 (1982).

25. Bryan, K. A. *Rev. Earth planet. Sci.* **10**, 15–38 (1982).
26. Hsiung, J. J. *phys. Oceanogr.* **15**, 1405–1413 (1985).
27. Broecker, W. S. *et al. Paleoceanography* **3**, 1–19 (1988).
28. Manabe, S. & Stouffer, R. J. *J. Clim.* **1**, 841–866 (1988).
29. Rooth, C. *Progr. Oceanogr.* **11**, 131–149 (1982).
30. Washington, W. M. & Mehl, G. A. *Clim. Dyn.* **4**, 1–38 (1989).
31. Dickson, R. R. *et al. Progr. Oceanogr.* **20**, 103–151 (1988).
32. Brewer, P. G. *et al. Science* **222**, 1237–1239 (1983).
33. Bradley, R. S. *et al. Science* **237**, 171–175 (1987).
34. COHMAP Members *Science* **241**, 1043–1052 (1988).
35. Hutsen, W. H. *Science* **207**, 64–66 (1980).
36. Street-Perrott, F. A. & Harrison, S. P. in *Paleoclimate Data and Modeling* (ed. Hecht, A. D.) 291–340 (Wiley, New York, 1985).
37. Mason, I. M., Rapley, C. G., Street-Perrott, F. A. & Guzowska, M. *Proc. EARSeL/ESA Symp. European Remote Sensing Opportunities* 235–241 (European Space Agency, Paris, 1985).
38. Street-Perrott, F. A. & Roberts, N. in *Variations in the Global Water Budget* (eds Street-Perrott, F. A., Beran, M. & Ratcliffe, R. A. S.) 331–345 (Reidel, Dordrecht, 1983).
39. Gillespie, R., Street-Perrott, F. A. & Switsur, R. *Nature* **306**, 680–683 (1983).
40. Gasse, F. thesis, Univ. Paris-VI (1975).
41. Fontes, J.-C. *et al. Z. Gletscher, Glazialgeol.* **21**, 89–96 (1985).
42. Butzer, K. W. *Phys. Geogr.* **1**, 42–58 (1980).
43. Owen, R. B. *et al. Nature* **298**, 523–529 (1982).
44. Richardson, J. L. & Dussinger, R. A. *Hydrobiologia* **143**, 167–174 (1986).
45. Haberyan, K. A. *Freshwater Biol.* **17**, 429–436 (1987).
46. Talbot, M. R. *et al. Palaeoecol. Afr.* **16**, 173–192 (1984).
47. Servant, M. & Servant, S. in *The Sahara and the Nile* (eds Williams, M. A. J. & Faure, H.) 133–162 (Balkema, Rotterdam, 1980).
48. Rognon, P. *Bull. Ass. geogr. Fr., Paris* **433**, 77–84 (1976).
49. Michel, P. *Palaeoecol. Afr.* **12**, 297–306 (1980).
50. Talbot, M. R. *Quat. Res.* **16**, 201–220 (1981).
51. Hamilton, A. C., Taylor, D. & Vogel, J. C. *Nature* **320**, 164–167 (1986).
52. Metcalfe, S. E., Street-Perrott, F. A., Perrott, R. A. & Harkness, D. D. in *Proc. 8th Int. Diatom Symp.* (ed. Ricard, M.) 471–482 (Koeltz, Koenigstein, 1986).
53. Flores Díaz, A. in *Tlapacoya: 35,000 Años de Historia del Lago de Chalco* (eds Lorenzo, J. L. & Mirambell, L.) 109–156 (Instituto Nacional de Antropología e Historia, Mexico City, 1986).
54. Metcalfe, S. E. thesis, Univ. Oxford (1985).
55. Bloomfield, K. & Valastro, S. Jr *Overseas Geol. Min. Res.* **46** (1977).
56. Lambert, W. in *Tlapacoya: 35,000 Años de Historia del Lago de Chalco* (eds Lorenzo, J. L. & Mirambell, L.) 77–100 (Instituto Nacional de Antropología e Historia, Mexico City, 1986).
57. Hastenrath, S. *J. Atmos. Sci.* **33**, 202–215 (1976).
58. Nicholson, S. E. *J. Clim. appl. Met.* **25**, 1365–1381 (1986).
59. Kutzbach, J. E. & Street-Perrott, F. A. *Nature* **317**, 130–134 (1985).
60. Mix, A. C., Ruddiman, W. F. & McIntyre, A. *Paleoceanography* **1**, 43–66 (1986).
61. Schnitker, D. A. *Rev. Earth planet. Sci.* **8**, 343–370 (1980).
62. Duplessy, J.-C., Moyes, J. & Pujol, C. *Nature* **286**, 479–482 (1980).
63. Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233–244 (1983).
64. Curry, W. B. *et al. Paleoceanography* **3**, 317–341 (1988).
65. Duplessy, J.-C. *et al. Paleoceanography* **3**, 343–360 (1988).
66. Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784–787 (1982).
67. Shackleton, N. J. in *The Fate of Fossil Fuel CO₂ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 401–427 (Plenum, New York, 1977).
68. Boyle, E. A. *Paleoceanography* **3**, 471–489 (1988).
69. Kroopnick, P. M. *Deep-Sea Res.* **32**, 57–84 (1985).
70. Boyle, E. A. & Keigwin, L. *Nature* **330**, 35–40 (1987).
71. Jansen, E. & Veum, T. *Nature* **343**, 612–616 (1990).
72. Stoker, M. S. *et al. J. quat. Sci.* **4**, 211–222 (1989).
73. Broecker, W. S. *et al. Nature* **341**, 318–321 (1989).
74. Ruddiman, W. F. in *North America and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) 463–478 (Geol. Soc. Am., Boulder, 1987).
75. Jones, G. A. & Keigwin, L. D. *Nature* **336**, 56–59 (1988).
76. Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145–214 (1981).
77. Teller, J. T. in *North America and Adjacent Oceans during the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) 39–69 (Geol. Soc. Am., Boulder, 1987).
78. Rind, D. *et al. Clim. Dyn.* **1**, 3–33 (1986).
79. Fairbanks, R. G. *Nature* **342**, 637–642 (1989).
80. Dansgaard, W., White, J. W. C. & Johnsen, S. J. *Nature* **339**, 532–534 (1989).
81. Spaulding, W. G. & Graumlich, L. J. *Nature* **320**, 441–444 (1986).
82. Dyke, A. S. & Prest, V. K. *Géogr. phys. quat.* **41**, 237–263 (1987).
83. Mayewski, P. A., Denton, G. H. & Hughes, T. J. in *The Last Great Ice Sheets* (eds Denton, G. H. & Hughes, T. J.) 67–178 (Wiley, Chichester, 1981).
84. Flohn, H. & Nicholson, S. *Palaeoecol. Afr.* **12**, 3–21 (1980).
85. Möner, N.-A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **19**, 63–85 (1976).
86. Overpeck, J. T. *et al. Nature* **338**, 553–557 (1989).
87. Oglesby, R. J., Maasch, K. A. & Saltzman, B. *Clim. Dyn.* **3**, 115–133 (1989).
88. Adam, D. P. *Quat. Res.* **5**, 161–171 (1975).
89. Williams, M. A. J. & Faure, H. (eds) *The Sahara and the Nile* (Balkema, Rotterdam, 1980).
90. Niederberger, C. *Science* **203**, 131–142 (1979).
91. Keigwin, L. D. & Jones, G. A. *Deep-Sea Res.* **36**, 845–867 (1989).

ACKNOWLEDGEMENTS. We thank S. E. Metcalfe for supplying her diatom counts and G. Delibrias, J.-C. Fontes, R. Gillespie, D. D. Harkness and R. Olet for radiocarbon dating. We also thank L. D. Keigwin, K. Bryan, C. K. Folland, S. Hastenrath, P. C. Lamb, A. C. Mix, R. E. Newell, W. F. Ruddiman and J. T. Teller for discussions. Our fieldwork in Africa and Mexico was supported by NERC, the Royal Society, the Royal Geographical Society and Oxford University.

ARTICLES

Evidence for two-step deglaciation and its impact on North Atlantic deep-water circulation

E. Jansen & T. Veum

Department of Geology, University of Bergen, Allégaten 41, N-5007 Bergen, Norway

Oxygen and carbon isotope records from benthic and planktonic foraminifera are presented for the past 35,000 years in the northeastern Atlantic. The results support the idea that the last deglaciation took place in two major steps^{1–4}, and conflict with theories calling for a strong reduction in North Atlantic deep-water formation to explain the abrupt cooling of the Younger Dryas cold period^{5–7}.

MAJOR deglaciations, which melted the Northern Hemisphere ice sheets over short time periods, characterize the last 600,000 years of global climate^{8,9}. It has long been known^{10–13} that in the last deglaciation, the temperature increase during the glacial meltdown was interrupted by a sharp return to almost full glacial temperatures in the North Atlantic and Europe (the Younger Dryas oscillation at 11,000–10,000 years before present (yr BP)). Whether the actual ice-sheet disintegration occurred in steps, accompanying this stepwise temperature trend, or whether the deglaciation was smooth and gradual, is still debated³. A convincing case for the stepwise deglacial model was made recently by Fairbanks⁴, who showed that the sea-level rise caused by the return of deglacial melt water to the ocean was characterized

by two periods of extremely rapid rise centred at 12,000 and 9,500 yr BP. The existence of abrupt climate shifts, occurring in less than a century, suggests strong nonlinearity in the response of the climate system to orbital forcing. Theories that explain such rapid climate shifts invoke abrupt reorganization of the surface and deep ocean circulation of the North Atlantic^{5,6,14}, either alone or in cooperation with shifting atmospheric flow patterns as the Laurentide ice sheet decreased³. As the present North Atlantic ocean circulation leads to large-scale meridional heat flux to the high northern latitudes and to deep-water formation and ventilation which is suggested to provide a controlling effect on variations in natural atmospheric CO₂ levels^{15,16}, a reorganization of ocean circulation in the past would have promoted large climate effects.

Chronology

To provide a detailed description of circulation changes affecting the North Atlantic during the deglaciation, we performed a stable-isotope study on a North Atlantic core V23-81, collected from the high-sedimentation-rate area of the Feni Drift west of Ireland (54°2' N, 16°8' W, at a water depth of 2,393 m). A detailed accelerator-mass-spectrometer radiocarbon chronology, covering the past 35,000 years, is available for this core^{7,17} (Table 1). Until now, no stable-isotope record from V23-81 was available,

TABLE 1 AMS dates from core V23-81

Depth (cm)	<i>N. pachyderma sin.</i>	<i>G. inflata</i>	Corrected age*
0-3		2,070 ± 90	1,670
7-8		1,820 ± 90	1,420
61-63		6,260 ± 150	5,860
135-136		9,490 ± 200	9,090
143-144	10,120 ± 180	10,450 ± 200	10,050
147-148		10,260 ± 190	9,860
151-158	11,300 ± 140		10,400
157-158	10,230 ± 200		
164-165		11,500 ± 200	11,100
171-172		10,530 ± 160	
172-173		10,960 ± 190	
175-176	10,780 ± 190 10,990 ± 190		
180-181		12,240 ± 220	
186-187		10,500 ± 230	
188-190	11,850 ± 200	11,330 ± 230	11,450
194-195	12,660 ± 240	11,940 ± 210	12,260
198-199	12,720 ± 220	12,530 ± 220	12,320
201-202		12,390 ± 240	
206-207		13,240 ± 310	12,840
215-216	14,060 ± 210		13,660
234-235	15,600 ± 190		15,200
293-294	17,140 ± 240		16,740
335-336	18,790 ± 280		18,390
384-385	24,820 ± 870		24,420
418-419	29,400 ± 960		29,000
449-450	32,540 ± 1240		32,140
499-500	35,640 ± 1810		35,240

Data are from ref. 17.

*Ages used to provide age model for Fig. 1, corrected for a reservoir age of 400 yr.

partly because of the low abundances of benthic foraminifers resulting from the high sedimentation rates (~15 cm per 1,000 yr). We produced detailed benthic and planktonic oxygen and carbon isotope records by pooling available sample material and using a high-sensitivity mass spectrometer (Table 2, Fig. 1). This gives a complete benthic and planktonic stable-isotope record for the North Atlantic directly dated by acceleration mass spectrometry (AMS). Duplessy *et al.*¹⁸ published AMS-dated records from core CH73-139C from the Feni Drift but a suggested hiatus in the first deglacial interval¹⁹ limits its value.

Large fluctuations in foraminiferal abundance lead to differences between AMS dates performed on different foraminiferal species from the same sample, even in high-sedimentation-rate cores where bioturbation effects are small²⁰. The species abundances are highly variable in V23-81, especially at the Bølling-Allerød/Younger Dryas/Holocene transition (13,000-9,000 yr BP). This causes some scatter in ages obtained from different species in the core⁷. To avoid bioturbation effects, we used the chronology of Broecker *et al.*¹⁷ based on the dates that were determined from the most abundant of the species *Neogloboquadrina pachyderma* or *Globorotalia inflata* (Table 1). Consistently higher ages with depth are recorded except in the Younger Dryas. This age reversal might be due to increased sedimentation rates, variations in ¹⁴C production or variations in the rate of ocean-atmosphere ¹⁴C transfer^{7,19}. The age for the Younger Dryas seems to be relatively accurate, based on the peak abundance of the polar planktonic foraminifer *N. pachyderma sin.* and a distinct peak of volcanic ash at 157 cm (Fig. 1), which has been correlated geochemically to the Vedde Ash layer²¹, dated²² in limnic sediments from western Norway at 10,600 ± 60 yr BP. It is AMS dated in V23-81 at 10,400 yr BP by the mid-point of two dates, 1,000 years apart¹⁷.

Two-step deglacial $\delta^{18}\text{O}$ signal

The benthic and planktonic $\delta^{18}\text{O}$ records (Fig. 1) show that the major deglacial $\delta^{18}\text{O}$ decrease started after 15,200 yr BP with

a first step lasting until 11,500 yr BP, and a second step at 10,000 yr BP. The general timing of these steps (Termination IA and IB, ref. 1) is in accordance with Fairbanks' sea-level record⁴, with Bard and collaborators¹⁹ AMS-dated planktonic record from off the coast of Portugal, with evidence^{7,23} of extensive melt-water drainage from the Laurentide ice sheet between 14,000 and 12,000 yr BP to the Gulf of Mexico, and with a low-salinity excursion at 14,500 yr BP documented²⁴ from the Fram Strait. Both the V23-81 and Bard and collaborators' planktonic record display a major $\delta^{18}\text{O}$ increase in the Younger Dryas. In V23-81, the benthic record shows that the second deglacial step (Termination IB) started between 10,400 and 10,000 yr BP. This phase is less well defined in the planktonic record, probably because of bioturbation. The abundance of left-coiling *N. pachyderma* decreases by orders of magnitude at ~10,000 yr BP, indicating that the trace amounts found above this level are contaminated by isotopically heavy specimens brought upwards by bioturbation. Bard *et al.*²⁰ documented and modelled identical effects in a study of the nearby core CH73-139C.

To what degree are the amplitude and shape of these isotope records determined by the deglacial transfer of light oxygen to the ocean, and to what degree are they determined by the temperature? The absolute benthic and planktonic glacial-to-Holocene $\delta^{18}\text{O}$ amplitudes are 1.9‰ and 2.75‰, respectively. Assuming²⁵ that 10 m of sea level is equivalent to 0.11‰ $\delta^{18}\text{O}$, estimates^{4,26,27} of the effect of global ice volume on glacial $\delta^{18}\text{O}$ are ~1.2‰ $\delta^{18}\text{O}$. The excess benthic $\delta^{18}\text{O}$ amplitude in V23-81 indicates that a strong temperature signal also contributes to the deglacial $\delta^{18}\text{O}$ shift in this record, and shows that glacial deep waters were ~3 °C colder than at present, in accordance with ref. 27. The even larger amplitude in the planktonic record indicates a glacial-to-Holocene temperature change of ~7 °C, or larger if glacial salinities were lower than at present.

The low $\delta^{18}\text{O}$ benthic peak ('overshoot') between 13,000 and 11,000 yr BP and the subsequent increase in the Younger Dryas cannot be explained as a temperature signal. During the 'overshoot', benthic $\delta^{18}\text{O}$ values were the same as in the Holocene. If this were a temperature signal, it would require that deep-water temperatures were 3 °C warmer than at present. This is unlikely at a water depth of 2,400 m, and we conclude that the benthic $\delta^{18}\text{O}$ 'overshoot' at Termination IA was caused by strong transfer of low $\delta^{18}\text{O}$ glacial melt water. Similar and larger benthic $\delta^{18}\text{O}$ 'overshoots' before and after the Younger Dryas are seen in Norwegian sea records (ref. 28 and T.V., unpublished data from Norwegian Sea core HM52-43). The Barbados sea-level curve⁴ shows that the sea level rose by ~10 m in the Younger Dryas, which must have caused a small global $\delta^{18}\text{O}$ decrease. The Younger Dryas benthic $\delta^{18}\text{O}$ increase thus reflects a reduced transfer of melt-water products, in close agreement with the Barbados sea-level record, which demonstrates that the melt-water discharge was reduced by 80% during the Younger Dryas. Most of the total glacial-to-Holocene benthic $\delta^{18}\text{O}$ decrease happened during Termination IA (Fig. 1), indicating that most of the deep-water temperature increase also took place in this part of the deglaciation, and that the Younger Dryas deep-water temperatures were similar to today's.

The benthic deglacial $\delta^{18}\text{O}$ 'overshoot', the $\delta^{18}\text{O}$ increase of the Younger Dryas and the next negative $\delta^{18}\text{O}$ shift at Termination IB strongly indicate that the deglaciation occurred in two discrete steps. The $\delta^{18}\text{O}$ steps in deep-sea benthic records are thus not artefacts of temperature overprints³, but instead reflect changing rates of deglaciation and melt-water transfer. The timing of the steps closely matches the two intervals of most rapid sea-level change documented in the Barbados sea-level record⁴ and corresponds with steps apparent in many palaeoclimatic records^{9,29-34}. Although some evidence indicates that the Laurentide ice sheet disintegrated in steps, this has not been clearly documented, and lack of firm evidence has led to the suggestion of a gradual deglaciation for this ice sheet³. The V23-81 $\delta^{18}\text{O}$ record indicates that the Laurentide ice sheet,

being a large part of the global ice-volume signal, experienced two phases of rapid disintegration, separated by a slower rate in the Younger Dryas.

In contrast to the benthic record, it is possible to explain the planktonic $\delta^{18}\text{O}$ 'overshoot' as an effect of increased temperature. Because the benthic 'overshoot' cannot be explained by temperature, however, the strong covariance between the benthic and planktonic $\delta^{18}\text{O}$ signals in V23-81 (Fig. 1) is an indication that the planktonic 'overshoot' on Termination IA might best be explained as a result of deglacial melt-water input. The timing indicates that it corresponds to the period immediately preceding the Younger Dryas. If this interpretation is correct, then the rerouting of deglacial melt water from the Mississippi to the St Lawrence during the Younger Dryas^{7,35} did not leave a marked impact on the North Atlantic $\delta^{18}\text{O}$ records, suggesting that the lowered salinity of North Atlantic surface waters was governed by the rate of deglaciation, rather than melt-water pathways. The possibility of different explanations for the benthic and planktonic $\delta^{18}\text{O}$ records can be resolved only by a detailed mapping of both $\delta^{18}\text{O}$ and palaeotemperature gradients by analysing a larger number of high-resolution cores in the North Atlantic with the aim of resolving both the geographical and seasonal distribution of the signals.

An interesting feature of the benthic $\delta^{18}\text{O}$ record from V23-81 is the high-amplitude fluctuation (1‰) during the period 32,000–18,000 yr BP, corresponding with upper oxygen isotope stage 3 and lower stage 2. This amplitude is greater than that of other benthic $\delta^{18}\text{O}$ records²⁶, indicating that the North Atlantic deep water warmed by 2 °C, or that the deep waters were influenced by melt water from a glacial retreat in upper isotope stage 3. Norwegian coastal records show that the North European ice sheet had retreated markedly at this time, before its final advance to its maximum position^{36,37}. The New Guinea sea-level record indicates²⁶ that sea level rose to –44 m at ~28,000 yr BP. If we assume that North Atlantic deep waters warmed by 2 °C, the $\delta^{18}\text{O}$ difference between the 24,000 yr BP level and the Holocene is in accord with the New Guinea sea-level data. We note, however, that today there is a 0.35‰ difference in $\delta^{18}\text{O}$ between North Atlantic Deep Water (NADW) and Antarctic Bottom Water (AABW)^{38,39}. It has been demonstrated^{40–43} that the production rate of NADW was reduced in glacials and in parts was replaced by AABW. A complete shift of source waters could produce a reduction in the $\delta^{18}\text{O}$ of benthic foraminifers comparable in size to a 1–2 °C temperature increase.

Transfer of the deglacial signal

The location of the core is close to potential deglacial melt-water sources, where the transfer of light oxygen from the melting ice sheets might be undiluted enough to create a significant local isotope signal, even in the deep ocean. Together with similar 'overshoots' in deep-sea benthic records from the Norwegian Sea (ref. 28 and T. V., unpublished data from Norwegian Sea core HM52-43), this indicates that there is a mechanism in the high northern latitudes to bring the light isotopically spiked water down to the deep ocean, despite the increased buoyancy caused by the fresh melt-water input. The $\delta^{18}\text{O}$ 'overshoot' in core V23-81 is of the order of 0.6–0.8‰, indicating a salinity decrease of at least 1‰ in the source areas. The general effect of such a reduction in surface salinity and density is to stabilize the water column and prevent the transfer of the deglacial signal to the deep-water realm through open-ocean vertical overturn. If, on the other hand, the deep waters formed by densification owing to sea-ice freezing processes, it is possible to maintain a light oxygen isotope signal while increasing salinity and density enough to form deep water. This process might be an important transport mechanism for the deglacial signal.

Deep-water formation and the Younger Dryas

The benthic $\delta^{13}\text{C}$ data in Fig. 1. have an amplitude of ~1‰. There are two prominent periods of low $\delta^{13}\text{C}$ with values close to 0‰, centred at 18,000 yr BP and 13,000 yr BP. Both the benthic and planktonic $\delta^{13}\text{C}$ records show positive $\delta^{13}\text{C}$ values in the Younger Dryas, followed by a more negative interval at 10,000 yr BP. The $\delta^{13}\text{C}$ of the deep Atlantic Ocean is generally thought to reflect the relative importance of nutrient-poor NADW (high $\delta^{13}\text{C}$, well oxygenated), and nutrient-enriched AABW (lower $\delta^{13}\text{C}$, less oxygenated)⁴⁴. Low benthic $\delta^{13}\text{C}$ therefore denotes periods with reduced NADW flux relative to AABW and vice versa. Low $\delta^{13}\text{C}$ values in the North Atlantic at 18,000 yr BP are well documented^{40–43} and are supported by high Cd/Ca levels⁴¹. The next low $\delta^{13}\text{C}$, located at the mid-point of Termination IA at 13,000 yr BP also most probably reflects a reduction of NADW production. Reduced surface-water salinities, because of enhanced deglacial melt-water flux during this time, caused increased buoyancy and reduced ventilation. The timing of this event is slightly earlier than the large melt-water discharge centred at 12,000 yr BP documented by the Barbados sea-level record⁴, and also ~1,000 years earlier than a low $\delta^{13}\text{C}$ level found⁴⁵ in a core from the Bermuda rise. It is, however, clearly

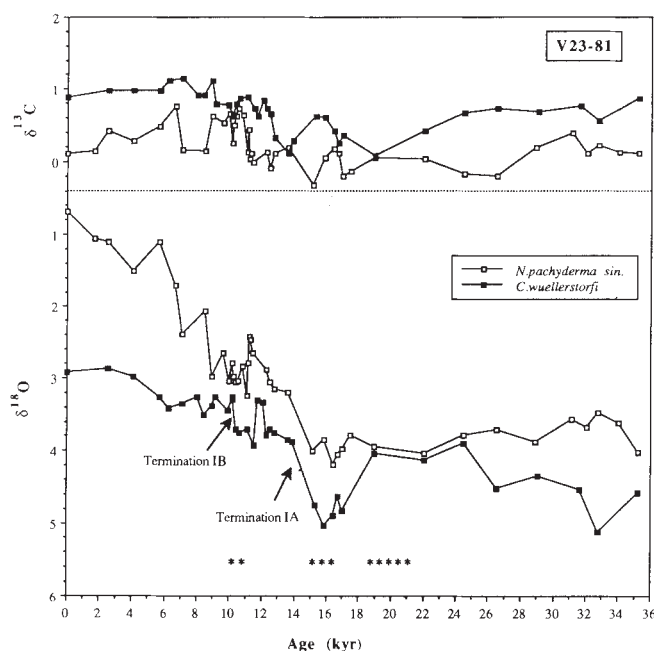


FIG. 1 Oxygen and carbon isotope records for the planktonic foraminifer *N. pachyderma sin.* and the benthic foraminifer *C. wuellerstorfi* from core V23-81 plotted against age. The age model is from ref. 17 and listed in Table 1. Asterisks indicate peak abundance levels of the polar planktonic foraminifer *N. pachyderma sin.*

TABLE 2 Stable-isotope data for core V23-81

Depth (cm)	<i>N. pachyderma sin.</i>		<i>C. wuellerstorfi</i>		Depth (cm)	<i>N. pachyderma sin.</i>		<i>C. wuellerstorfi</i>	
	$\delta^{13}\text{C}$ (‰)	$\delta^{18}\text{O}$ (‰)	$\delta^{13}\text{C}$ (‰)	$\delta^{18}\text{O}$ (‰)		$\delta^{13}\text{C}$ (‰)	$\delta^{18}\text{O}$ (‰)	$\delta^{13}\text{C}$ (‰)	$\delta^{18}\text{O}$ (‰)
0	0.10	0.69	0.88	2.92	180	0.03	2.47	0.62	3.31
10	0.14	1.06			186	0.11	2.66		
20	0.41	1.11	0.97	2.87	189			0.83	3.34
40	0.28	1.51	0.98	2.99	198	-0.02	2.90	0.72	3.79
60	0.47	1.11	0.98	3.26	201	0.12	3.06	0.65	3.72
70			1.12	3.42	206	-0.09	3.15	0.32	3.77
80	0.75	1.71			215	0.10	3.21	0.10	3.86
90	0.15	2.40	1.14	3.36	218			0.28	3.89
110			0.92	3.27	234	0.18	4.01		
120			0.92	3.51	237			0.61	4.75
122	0.13	2.07			260	-0.33	3.86	0.60	5.03
131	0.62	2.99	1.12	3.39	280	0.05	4.20	0.42	4.89
136			0.79	3.27	293	0.17	4.06	0.25	4.64
140	0.37	2.66			300	0.10	3.98	0.36	4.83
142			0.78	3.45	313	-0.20	3.79		
143	0.52	3.05			340	-0.15	3.95	0.08	4.05
147			0.64	3.30	365	0.05	4.04	0.42	4.14
148	0.65	2.80	0.62	3.27	385	0.02	3.79	0.67	3.91
151	0.24	2.98			400	-0.17	3.71	0.73	4.53
154	0.49	3.07	0.79	3.71	418	-0.20	3.88	0.68	4.36
157	0.62	3.05			440	0.19	3.57		
158			0.86	3.76	445			0.75	4.54
161	0.73	2.85			449	0.38	3.68		
164			0.88	3.71	460	0.11	3.48	0.56	5.13
166	0.64	3.25			480	0.21	3.62		
170	0.12	2.80			499	0.12	4.03	0.87	4.58
175	0.43	2.43	0.73	3.94					

$\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}] - 1$, where the standard used was PDB belemnite. The definition of $\delta^{18}\text{O}$ is analogous; the PDB belemnite standard was also used.

within the error limits of the timescales for these records and the V23-81 record, and is probably a correlated event. It thus reflects a large reduction in NADW production and a replacement of NADW by AABW triggered by the first big release of deglacial melt water to the ocean. The benthic low $\delta^{13}\text{C}$ level at 13,000 yr BP coincides with light planktonic $\delta^{18}\text{O}$ values, indicative of decreased surface salinities. This low $\delta^{18}\text{O}$ is probably under-represented by the data because of bioturbation. At this level the abundance of *N. pachyderma sin.* drops markedly, which makes it likely that the $\delta^{18}\text{O}$ values are made heavier by contamination from lower and more abundant levels. This is supported by preliminary isotope analyses of the more abundant species *N. pachyderma dex.*, which record a much larger $\delta^{18}\text{O}$ depletion at this level (E. J. and M. Odland, unpublished data).

The $\delta^{13}\text{C}$ rise into the Younger Dryas indicates that NADW formation and deep-water ventilation increased, probably as a response to a reduction in the total melt-water release to the critical areas of deep overturn. The small minima at 11,500 and 10,000 yr BP might reflect minor reductions in the flux of ventilated deep water at V23-81. Heavy $\delta^{13}\text{C}$ values in the Younger Dryas, bracketed by $\delta^{13}\text{C}$ minima at 12,000–13,000 yr BP and 10,000–9,000 yr BP, are also clearly seen in Norwegian Sea benthic and planktonic records (ref. 28 and T.V., unpublished data from Norwegian sea core HM52-43), and in planktonic and benthic deglacial $\delta^{13}\text{C}$ records from the Caribbean^{42,46}. The data also agrees with the record⁴⁷ from core CH73-139C from the same areas as V23-81. From this evidence it seems that, approaching the Younger Dryas, there was a resumption of vertical convection and deep-water ventilation after a period of large influx of low-salinity melt water during Termination IA. Consequently, the Younger Dryas cold phase probably did not occur as a response to a cessation of NADW formation caused by a re-routing of melt-water drainage from the Mississippi to the St Lawrence, as suggested by Broecker *et al.*⁵, because this would require a reduction of $\delta^{13}\text{C}$ values at the initiation of the Younger Dryas.

Our evidence for NADW production during the Younger Dryas is not in agreement with the conclusions of Boyle and

Keigwin⁴⁵ who, based on $\delta^{13}\text{C}$ and Cd records from core EN120GGC1 from the Bermuda Rise, argued that NADW was markedly reduced during the Younger Dryas. Their evidence has been taken as a confirmation of the melt-water-diversion model for explaining the Younger Dryas^{7,35}. The Younger Dryas level is well established in V23-81, because it is directly tied by ash-layer correlation to the continental Younger Dryas cold phase²¹. As pointed out by Fairbanks, the minimum $\delta^{13}\text{C}$ level in the Bermuda Rise record can easily be interpreted to occur at the termination of the Younger Dryas at ~10,000 yr BP, and not at the initiation (see Fig. 5 in ref. 4). Consequently, with only a slight adjustment of the timescale used by Boyle and Keigwin, the Bermuda Rise record is consistent with the V23-81 data.

If the EN120GGC1 timescale is correct, there are ways to explain the apparent contradiction between the Cd/Ca record of the Bermuda Rise and the $\delta^{13}\text{C}$ record of V23-81. One possibility is that $\delta^{13}\text{C}$ records do not only reflect NADW production fluctuations. As the $\delta^{13}\text{C}$ and the nutrient content (and Cd content) might be decoupled in the surface waters of the source regions of NADW⁴⁸, it is possible that $\delta^{13}\text{C}$ variations may merely record the preformed $\delta^{13}\text{C}$ of the source waters rather than ventilation rate, respiration and deep-water formation, which would be recorded by Cd/Ca ratios. The low $\delta^{13}\text{C}$ levels in the V23-81 record may therefore not only reflect a cessation of NADW production. It is much more difficult, however, to reconcile the high $\delta^{13}\text{C}$ values in V23-81 and high Cd/Ca at the Bermuda Rise in the Younger Dryas, because high $\delta^{13}\text{C}$ values reflect well oxygenated and ventilated deep waters and high Cd/Ca ratios reflect poorly ventilated waters. Another possibility is that the Bermuda Rise site and V23-81 were at times located in different water masses. V23-81 is located at a site where the present deep waters consist of a mixture of northeast Atlantic deep water with a small AABW component, and Norwegian Sea overflow water. As the main deep-water overflow today passes through the Denmark Strait (west of V23-81), V23-81 does not record the chemical composition of the Denmark Strait overflow. The combined mixture of the

different overflows and other components of NADW would only be seen further downstream in the NADW, where EN120GGC1 is located. The Bermuda Rise core and V23-81 are also from different depths (4,450 m and 2,393 m respectively). Discrepancies between the records may therefore be due to exceptionally steep chemical gradients in the deep water, reflecting a shallower and less extensive NADW than at present. Strong deep-water $\delta^{13}\text{C}$ gradients within the North Atlantic during glacials were recently described by Raymo *et al.*⁴⁹. The high $\delta^{13}\text{C}$ at the Younger Dryas level in V23-81 indicates, however, that the formation of NADW, although possibly more limited than today, was more like the present circulation than that of the last glacial maximum.

Although it is possible to construct scenarios that explain both the V23-81 and the Bermuda Rise records, the discrepancy is most easily explained by timescale problems with the Bermuda Rise record. The available data clearly shows that the Norwegian Sea and the northeast Atlantic deep waters experienced a $\delta^{13}\text{C}$ rise approaching the Younger Dryas, and that at present there is no evidence to support hypotheses that deep-water ventilation ceased or was markedly reduced in the present source areas of NADW during the Younger Dryas. There is strong evidence that it was enhanced compared to the preceding and following intervals. The total deep-water production rate, however, which is the critical factor for the global impact of NADW, cannot be accurately determined by these data. We hope that present and planned activities will provide sufficient amounts of high-resolution large-diameter cores with enough sediment material to produce the AMS-dated records of both $\delta^{13}\text{C}$ and Cd/Ca needed to reconstruct accurately circulation changes in both the vertical and horizontal planes.

Implications

The isotope records from V23-81 place some constraints on models of climate change for the last deglaciation. The documentation of two deglacial steps suggests that the melting of ice-age ice sheets has a nonlinear response to orbital forcing. Palaeotemperature records and general circulation model (GCM) runs indicate^{3,50-52} that the high elevation of the Laurentide ice sheet cooled the ocean by advecting cold air over the North Atlantic. The deglacial reduction of ice-sheet height prob-

ably reduced this effect and led to a warmer ocean. GCM experiments indicate that ocean heating might explain the warm European during the Bølling-Allerød chronozones and the rapid deglaciation during the 13,000-12,000 yr BP interval. The cause of the steps, the reduced deglaciation rate and the sharp decrease of temperatures in the Younger Dryas remains an enigma and points to the existence of strong negative feedback mechanisms, which over a short time interval counteracted the orbital forcing. It is plausible that during the Younger Dryas, despite the cold climates of this period, there was in the northeast North Atlantic and Norwegian Sea a largely open, ice-free ocean with deep-water convection. Glaciological evidence from northwest Europe also indicates a relatively sea-ice-free ocean. Such a situation is needed to provide enough moisture to obtain the observed lowering of glaciation limits in coastal areas in the Younger Dryas⁵³. Most evidence points to the periods before and after the Younger Dryas as the intervals that experienced the highest rate of melt-water transfer to the North Atlantic and the Norwegian Sea. Thus, melt-water discharge might be an important element in the climate feedbacks that shaped the deglaciation, but there is at present no firm evidence from the present regions of NADW formation that the Younger Dryas was caused by a melt-water anomaly.

The most likely chain of events indicated by the V23-81 records is as follows. The first deglacial step led to a sharp reduction in vertical overturn and formation of northern component water. Increasing surface ocean temperature in the Bølling-Allerød chronozones led to increased continental temperatures that provided a positive feedback for a further enhancement of ice-sheet disintegration. Approaching the Younger Dryas, deep-ocean ventilation commenced, probably because of a combination of surface salinity changes and changes in the wind field along the marginal ice zone^{54,55}. The Younger Dryas was characterized by both reduced deglaciation rates and deglacial melt-water release and colder oceanic and continental temperatures in Europe, in a setting with a relatively open and ice-free ocean. The causes of the Younger Dryas cooling are still unclear. Shifts in atmospheric circulation caused by albedo feedbacks from ice-sheet disintegration and shifting ice-sheet morphology, or temperature shifts caused by sudden drops in atmospheric CO_2 , should be further explored. □

Received 4 August 1989; accepted 4 January 1990.

- Duplessy, J. C., Delibrias, G., Turon, J. L., Pujol, C. & Duprat, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 121-144 (1981).
- Mix, A. C. & Ruddiman, W. F. *Quat. Sci. Rev.* **4**, 59-108 (1985).
- Ruddiman, W. F. in *North America and Adjacent Oceans During the Last Deglaciation* Chs 7, 21 (eds Ruddiman, W. F. & Wright, H. E. Jr) (Geol. Soc. Am., Boulder, 1987).
- Fairbanks, R. G. *Nature* **342**, 637-642 (1989).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21-26 (1985).
- Rind, D. *et al. Clim. Dyn.* **1**, 3-33 (1986).
- Broecker, W. S. *et al. Palaeoceanography* **3**, 1-19 (1988).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
- Ruddiman, W. F., Raymo, M. & McIntyre, A. *Earth planet. Sci. Lett.* **80**, 117-129 (1986).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145-214 (1981).
- Mangerud, J. *Norsk. geol. Tidsskr.* **57**, 23-54 (1977).
- Oeschger, H. *et al. Am. geophys. Un., Geophys. Monogr.* **29**, 299-306 (1984).
- Overpeck, J. T. *et al. Nature* **338**, 553-557 (1989).
- Berger, W. H. & Vincent, E. *Nature* **324**, 53-55 (1986).
- Boyle, E. A. *J. geophys. Res.* **93**, 15701-15714 (1988).
- Broecker, W. S. & Peng, T.-H. *Global Biogeochemical Cycles* (1989).
- Broecker, W. S. *et al. Quat. Res.* **30**, 1-6 (1988).
- Duplessy, J. C. *et al. Nature* **320**, 350-352 (1986).
- Bard, E. *et al. Nature* **328**, 791-794 (1987).
- Bard, E. *et al. Clim. Dyn.* **1**, 101-112 (1987).
- Kvamme, T., Mangerud, J., Furnes, H. & Ruddiman, W. F. *Norsk. geol. Tidsskr.* (in the press).
- Mangerud, J. *et al. Quat. Res.* **21**, 85-104 (1980).
- Broecker, W. S. *et al. Nature* **341**, 318-329 (1989).
- Jones, G. A. & Keigwin, L. D. *Nature* **336**, 56-59 (1988).
- Fairbanks, R. G. & Mathews, R. K. *Quat. Res.* **10**, 181-196 (1978).
- Chappell, J. & Shackleton, N. J. *Nature* **324**, 137-140 (1986).
- Labeyrie, L. D. *et al. Nature* **327**, 477-482 (1987).
- Jansen, E. & Erlenkeuser, H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **49**, 189-206 (1985).
- Mangerud, J. in *Studies in the Late Glacial of Northwest Europe* (eds Gray, J. M. & Robinson, J. M.) 22-30 (Pergamon, Oxford, 1980).

- Van der Hammen, T. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **32**, 247-340 (1981).
- Jakobsen, G. L. Jr, Webb, T. III & Grimm, E. C. in *North America and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) 277-288 (Geol. Soc. Am., Boulder, 1987).
- Dansgaard, W. *et al. Am. geophys. Un., Geophys. Monogr.* **33**, 71-76 (1985).
- Siegenthaler, U., Eicher, U. & Oeschger, H. *Ann. Glaciol.* **5**, 149-152 (1984).
- Vorren, T. O. *et al. Boreas* **17**, 41-77 (1988).
- Broecker, W. S. *et al. Nature* **341**, 318-321 (1989).
- Mangerud, J. *Palaeoklimatologie* (Gustaf Fisher, Stuttgart, 1988).
- Larsen, E. *et al. Boreas* **16**, 267-293 (1987).
- Craig, H. & Gordon, L. in *Stable Isotopes in Oceanographic Studies and Paleotemperatures* (ed. Tongiorgi, E.) 9-130 (CNR, Pisa, 1965).
- Jakobs, S. S., Fairbanks, R. G. & Horito, Y. *Am. geophys. Un., Antarctic Res. Ser.* **43**, 59-85 (1985).
- Mix, A. C. & Fairbanks, R. G. *Earth planet. Sci. Lett.* **73**, 231-243 (1985).
- Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784-787 (1982).
- Oppo, D. W. & Fairbanks, R. G. *Earth planet. Sci. Lett.* **86**, 1-15 (1987).
- Duplessy, J. C. *et al. Palaeoceanography* **3**, 343-360 (1988).
- Kroonnick, P. *Deep Sea Res.* **32**, 57-84 (1985).
- Boyle, E. A. & Keigwin, L. D. *Nature* **330**, 35-40 (1987).
- Oppo, D. W. & Fairbanks, R. G. *Palaeoceanography* **4**, 333-352 (1989).
- Labeyrie, L. D. & Duplessy, J. C. *Palaeoclimatol. Palaeogeogr. Palaeoecol.* **50**, 217-240 (1985).
- Johannessen, T. thesis, Univ. of Bergen (1987).
- Raymo, M. E. *et al. Palaeoceanography* **4**, 413-446 (1989).
- Kutzbach, J. E. & Guetter, P. J. *J. Atmos. Sci.* **43**, 1726-1759 (1986).
- Manabe, S. & Broccoli, A. J. *J. geophys. Res.* **90**, 2167-2190 (1985).
- Ruddiman, W. F. & McIntyre, A. *Geol. Soc. Am. Bull.* **95**, 381-396 (1984).
- Larsen, E., Eide, F., Longva, O. & Mangerud, J. *Arct. Alp. Res.* **16**, 137-160 (1984).
- Johannessen, J. A. *et al. J. geophys. Res.* **92**, 6754-6772 (1987).
- Crowley, T. J. & Häkkinen, S. *Palaeoceanography* **3**, 249-258 (1988).

ACKNOWLEDGEMENTS. We thank W. Broecker, W. Ruddiman and the Lamont-Doherty core repository for providing samples for this study. We also thank C. Ravelo, E. Bard, W. Broecker, R. Fairbanks, D. Oppo, L. Labeyrie, L. Keigwin, W. Ruddiman and H. Schrader for discussions and reviews, and O. Hansen for mass spectrometer operation. This study is a part of the project *Predicting Ocean Climate*, supported by the Norwegian Research Council for Science and the Humanities.

Cloning of the *T* gene required in mesoderm formation in the mouse

Bernhard G. Herrmann^{*†}, Siegfried Labeit[‡], Annemarie Poustka[§], Thomas R. King^{†||} & Hans Lehrach[¶]

^{*}Laboratory of Molecular Embryology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

[‡]European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG

[§]German Cancer Research Center, Im Neuenheimer Feld 280, 6900 Heidelberg, FRG

^{||}Laboratory of Genetics and McArdle laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706, USA

[¶]Imperial Cancer Research Fund Laboratory, Lincoln's Inn Fields, PO Box No. 123, London WC2A 3PX, UK

The murine developmental mutation *T* identifies an essential gene in mesoderm formation. Embryos lacking normal gene activity fail to form the notochord, the entire posterior region and the allantois, and die at about 10 days of gestation. We have isolated the *T* gene using a combination of molecular and genetic techniques, thus making molecular tools available to study processes underlying mesoderm formation in the mouse.

MANY murine developmental mutations have been investigated by genetic and embryological techniques, among them several loci required in mesoderm formation¹. But so far, genes identified only by mutations have not generally been accessible to molecular analysis.

We decided to use genetic and molecular techniques to isolate the *T* gene (*Brachyury*) located in the *T/t* complex on murine chromosome 17. Many mutant alleles of *T* have been identified and several have been studied (see refs 1 and 2 for reviews). Loss of function of the *T* gene leads to a disturbance of the primitive streak^{3,4} so that gastrulating *T/T* embryos generate insufficient mesoderm, whereas the number of ectodermal cells is increased⁵. The chordamesoderm is most strongly affected, and although the notochordal plate is formed initially, it later degenerates and no notochord is established. The posterior region of the embryo is entirely missing in late stages, probably owing to a failure of primitive streak regression. Finally, the allantois, a derivative of the mesoderm, is not formed, resulting in embryonic death at ~10 days of gestation⁶.

Heterozygotes for null alleles of *T* have a mild phenotype because of haplo-insufficiency. They have a short tail⁷ and tend to have one or a few malformed sacral vertebrae⁸. The defect is enhanced by the *tct* factor of *t* haplotypes (*T/t* animals are tailless); *tct* most probably is an allele of *T* with less than normal activity⁹. Mice carrying the gain-of-function allele *T^c* are more strongly affected than those carrying a deletion such as *T^{hp}*. *T^c/+* heterozygotes have severe malformations of the axial skeleton, rib fusions, absence of nuclei pulposi and lack of the tail¹⁰.

We describe here the cloning of the *T* gene by a deterministic approach, from the phenotype to the gene, using a combination of genetic and molecular techniques. We determined the fine structure of a 2,000 kilobase (kb) chromosomal region, cloned parts of it and identified the *T* gene by analysing several *T* alleles. *In situ* hybridization data obtained by Wilkinson *et al.*¹¹ show that the expression of the *T* gene occurs in presumptive and early stage mesoderm and becomes restricted to the notochord, tissues affected by the *T* mutation. The combined

data indicate that the *T* gene encodes a novel protein required in the formation of the mesoderm and in the morphogenesis of the notochord. The molecular analysis of *T* should provide insight into these fundamental processes of mammalian embryogenesis.

Positioning of the *T* gene

The cloning of a gene starting from the mapped position of a mutation is a multistep process. First, to obtain DNA probes located close to the gene, we have cloned chromosomal pieces dissected from spread metaphase chromosomes¹². Subsequent assignment of several DNA fragments to subregions of the *T/t* complex on chromosome 17 by genetic mapping identified two markers, 119 and 66, located near the *T* gene¹³. We then analysed the fine structure of a 1,400-kb chromosome piece marked by these two DNA fragments¹⁴. The results led to the identification of a large inverted duplication in the region, the establishment of the chromosomal orientation of two different cosmid clusters isolated by hybridization with the 119 clone (119I and 119II), and the localization of the centromeric limit of the deletion of the deficiency allele *T^{hp}* (Fig. 1a, b). Because *T^{hp}* is a null allele of *T*, *T* must be located within or at the break of the *T^{hp}* deletion¹⁵. These data indicate that 119II is closer to *T* than is 119I.

To establish the genetic distance between the cosmid cluster 119II and *T* with high resolution, we performed a genetic cross involving *T*, with the centromere (proximally) and *tf* (distally) as flanking markers (B.G.H. and H.L., unpublished work). Probes C and A (Fig. 1c) were tested, but neither probe was separated from *T* in 380 chromosomes scored. The results demonstrated very close genetic linkage between 119II and *T*.

The analysis of the allele *t^{Tu3}* identified the position of *T* relative to 119II. Allele *t^{Tu3}* is a chromosomal variant that complements the phenotype of the *T* mutation (*T/t^{Tu3}* animals have normal tails, whereas *T/+* animals have short tails)¹⁶. Allele *t^{Tu3}* contains a large duplication of part of the *T/t* complex including the *T* gene and the marker *Tcp1* (ref. 17). We have obtained results giving similar conclusions. In addition, we have localized the proximal end of the duplicated region between probe A and probe C. These data imply that the *T* gene is distal but close to probe C of the cosmid cluster 119II (Fig. 1b, c; data not shown).

Chromosome walking and jumping

The genetic analysis identified 119II as the cloned region closest to *T*, established its orientation on the chromosome and its position relative to *T*. 119II served as the starting point for the isolation of additional markers needed either to delimit the region containing the mutation, or to identify directly deletions or insertions that show the position of the gene in one or more *T* alleles.

We used a combination of chromosome walking and jumping¹⁸⁻²⁰ to isolate probes throughout a region extending 400 kb distal to probe III, the original DNA probe derived by

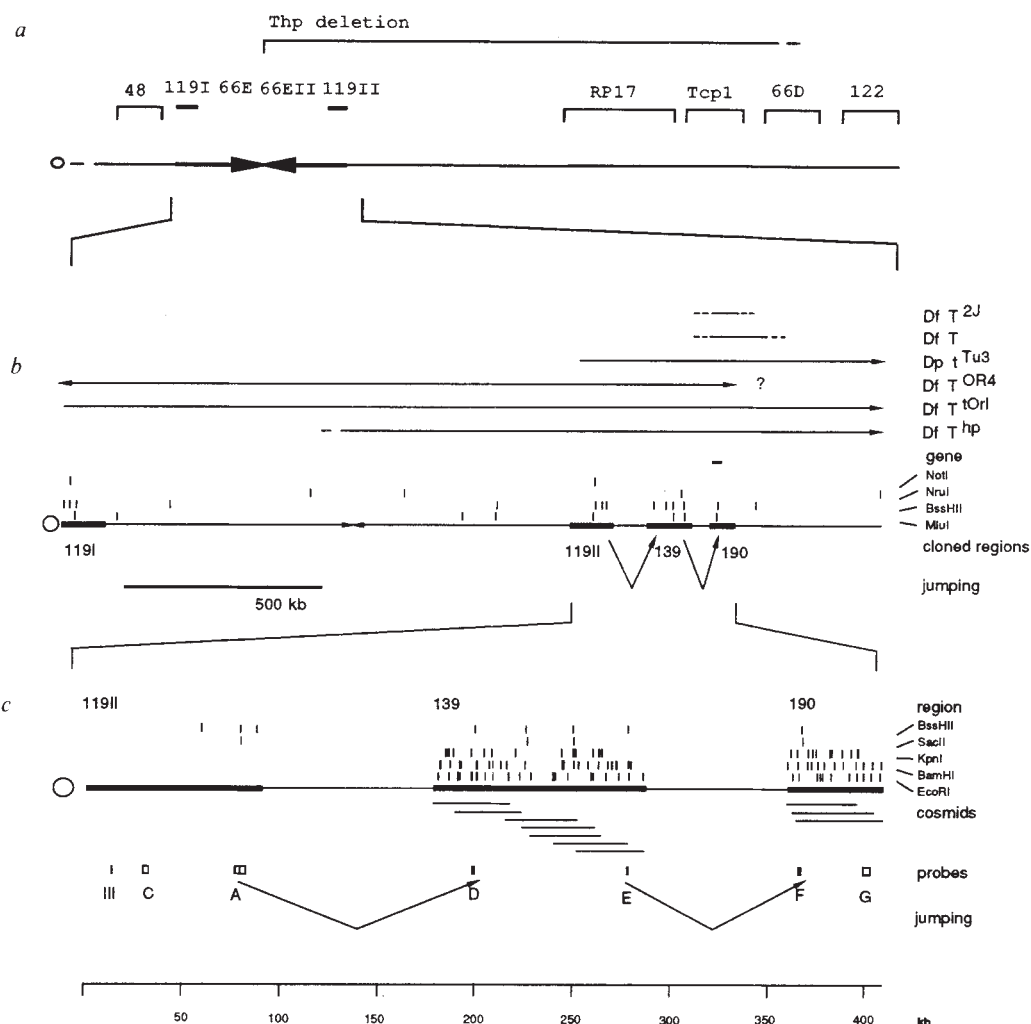
[†]Present addresses: Max Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35/II, 7400 Tübingen, FRG (B.G.H.) and The Jackson Laboratory, Bar Harbor, Maine 04609, USA (T.R.K.).

microdissection (Fig. 1c)¹⁴. Chromosome jumping, a technique that allows the cloning of the ends of DNA fragments of hundreds of kilobases is especially useful, because its directionality and large step size allows rapid progress through a chromosomal region, without the need to clone all intervening DNA sequences¹⁸. In addition, clones isolated from jumping libraries constructed with enzymes cutting preferentially in CpG islands (rare-cutter libraries) often mark the 5' ends of genes²¹, and therefore provide immediate access to candidate genes.

Identification of the *T* gene

The analysis of several *T* alleles with probe F (Fig. 1c) identified small deletions in the original isolate *T* and in *T*^{2J} (Fig. 2a). The mutant chromosomes contain altered restriction fragments that allowed the calculation of the sizes of both deletions by pulsed-field gel electrophoresis (PFGE; Fig. 2b)²². Genomic regions of 160–200 kb and 80–110 kb are deleted from the *T*- and *T*^{2J}-bearing chromosomes, respectively, localizing the *T* mutation to an interval of at most 110 kb, including probes F

FIG. 1 Genetic and molecular maps of the proximal part of chromosome 17 and the region near the *T* gene shown at three levels of resolution. **a**, Overview over the entire region analysed. **b**, Inverted duplication (119I–119II) and adjacent regions are shown at the resolution of **a** (rare-cutter enzyme) restriction map providing details on the extent of deficiency (Df) and duplication (Dp) alleles of the *T* mutation. Restriction sites are represented by vertical lines. The start and end of the deletions of *T* and *T*^{2J} were not precisely located, and are indicated by broken lines. Arrow heads indicate the continuation of the deficiencies or duplication; ?, telomeric extension of the deficiency beyond region 190 is uncertain. Jumping clones are symbolized by a V with an arrow-head facing the end-point fragment. Jumps start and end in *Bss*III–*Eco*RI fragments isolated in phage clones. **c**, Cloned regions, probes, isolated cosmids and jumping clones, and a restriction map of the cloned regions 139 and 190. The restriction map of 119II has been published previously¹⁴. Genetic linkage of 119II with *T* at high resolution was tested. For crossover in the interval centromere–*T*, 380 chromosomes were scored, and 808 were scored for recombination in the interval *T*–*tufted* (*tf*). The *tf* marker is located 7.5 centimorgans (cM; 1 cM is equivalent to a 1% recombination between two loci and corresponds to $\sim 2 \times 10^6$ base pairs of DNA) distal to *T* on chromosome 17. DNA samples of all recombinants in the intervals centromere–*T* (4) and *T*–*tf* (58) were analysed by Southern-blot hybridization. Probes A and C were not separated from *T* in 380 chromosomes scored. In addition probe C detected a polymorphic restriction fragment located in the region 119II that is always found in linkage with the original *T* mutation (data not shown). Both sets of data provided evidence for close genetic linkage of 119II with *T*. The analysis of the duplication allele *t*^{tu3} established the position of *T* distal to probe C of the 119II region (see text, and Fig. 1b, c). Chromosome jumping clones were isolated from a *Bss*III jumping library constructed from 129Sv DNA²⁰. Clones were verified by PFGE (ref. 22) and genetic mapping of the end-point fragments to exclude products of rare intermolecular ligation events. The first jumping clone isolated, connecting markers A and D, bridged a genomic fragment of ~ 120 kb. Marker D, the jumping end point, was then used to isolate a cluster of cosmids. Because the orientation of marker D on the



chromosome follows from the orientation of marker A, the isolated cosmids are immediately oriented. After a chromosome walk of 100 kb, performed to cross a region of several closely spaced *Bss*III sites, a second set of jumping clones was isolated. The end point of this jump (marker F) was again used to isolate a cluster of cosmids covering a total of 50 kb of DNA. The full designation of the DNA markers 139 and 190 should be D17Her139 and D17Her190, respectively, according to the rules for nomenclature. METHODS. The DNA of wild-type strains and *T* mutants was analysed by Southern blotting and hybridization according to published techniques^{13,32}. The CHEF system³³ was used for PFGE analysis¹⁴. For chromosome walking, whole cosmids (10–20 ng) were labelled³⁴, hybridized against total mouse DNA (100 µg) and vector DNA (2 µg) in (200 µl) solution (120 mM sodium phosphate, pH 6.8) for 2 h at 68 °C, and hybridized to cosmid DNA released from *Escherichia coli* colonies and bound to nylon filters by a modification of standard protocols¹⁴. Similarly, total cosmids of the cloned region 139 were used as probes on Southern blots in genetic analyses³⁵ and for the screening of the phage jumping library constructed as previously described²⁰ using the rare-cutter enzyme *Bss*III and *Eco*RI. Cosmid restriction maps were derived as described elsewhere^{36,37}.

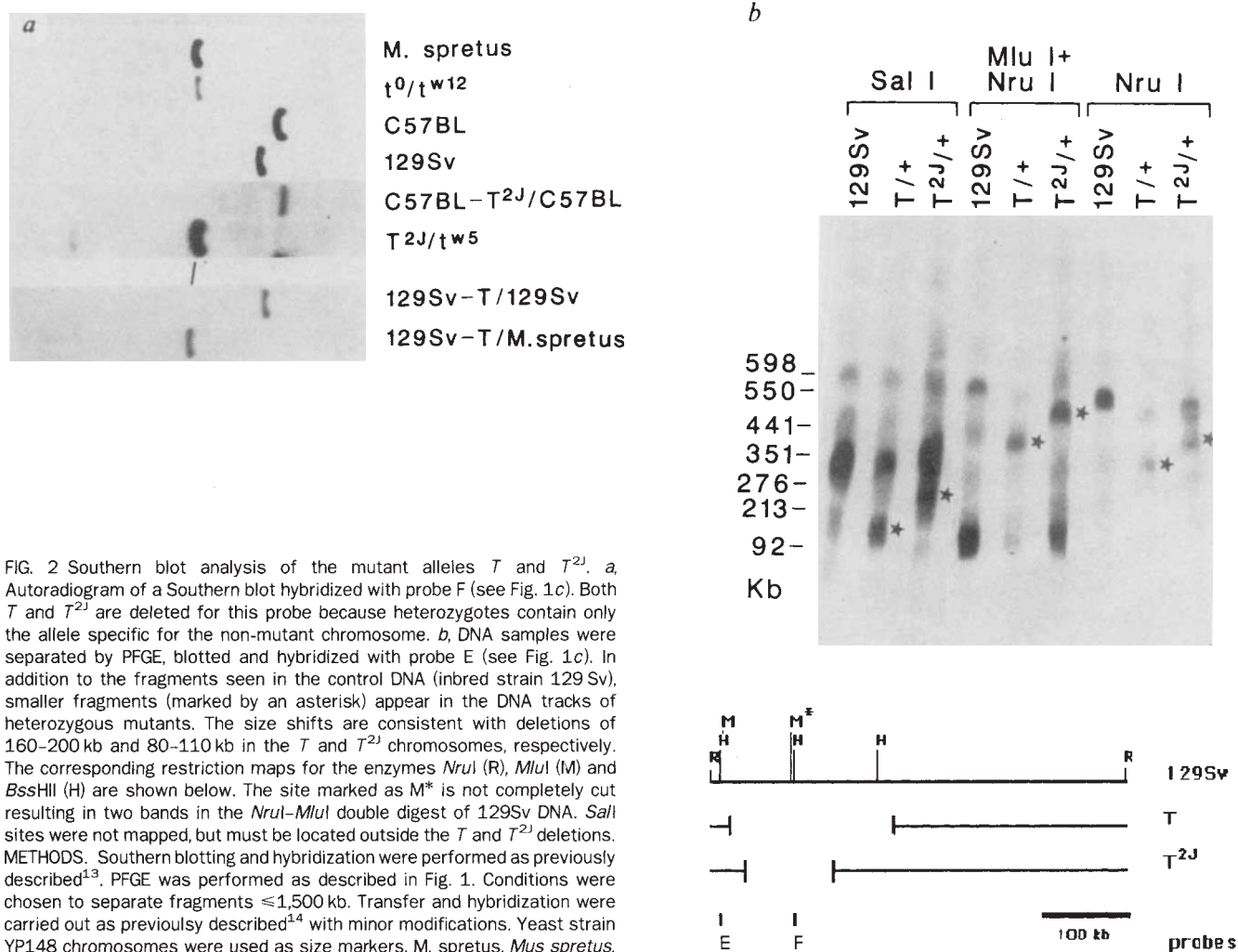


FIG. 2 Southern blot analysis of the mutant alleles *T* and *T*^{2J}. *a*, Autoradiogram of a Southern blot hybridized with probe F (see Fig. 1c). Both *T* and *T*^{2J} are deleted for this probe because heterozygotes contain only the allele specific for the non-mutant chromosome. *b*, DNA samples were separated by PFGE, blotted and hybridized with probe E (see Fig. 1c). In addition to the fragments seen in the control DNA (inbred strain 129Sv), smaller fragments (marked by an asterisk) appear in the DNA tracks of heterozygous mutants. The size shifts are consistent with deletions of 160–200 kb and 80–110 kb in the *T* and *T*^{2J} chromosomes, respectively. The corresponding restriction maps for the enzymes *Nru*I (R), *Mlu*I (M) and *Bss*HI (H) are shown below. The site marked as M* is not completely cut resulting in two bands in the *Nru*I–*Mlu*I double digest of 129Sv DNA. *Sal*I sites were not mapped, but must be located outside the *T* and *T*^{2J} deletions. METHODS. Southern blotting and hybridization were performed as previously described¹³. PFGE was performed as described in Fig. 1. Conditions were chosen to separate fragments $\leq 1,500$ kb. Transfer and hybridization were carried out as previously described¹⁴ with minor modifications. Yeast strain YP148 chromosomes were used as size markers. *M. spretus*, *Mus spretus*.

and G (data not shown) of the cosmid cluster 190.

This region contains one restriction site each for the rare-cutter enzymes *Bss*HI and *Sac*II, sites often found in CpG islands associated with genes²¹. We therefore screened this region for coding sequences expressed at 8.5 days of embryogenesis by hybridization of a cosmid to a complementary DNA library prepared from RNA of 8.5-day-old embryos²³. At this time of development, embryos are forming mesoderm and the notochord, tissues affected by the *T* mutation, and are probably expressing the *T* gene messenger RNA^{3,4}. We isolated several cDNA clones and analysed one, *pme75*, in detail. As expected, fragments detected by *pme75* are deleted from the deficiency alleles *T*, *T*^{2J}, *T*^{Or4}, *T*^{Or1} and *T*^{hp}, and duplicated in *t*^{Tu3} (Figs. 1b and 3a)^{1,17}. Among several more *T* alleles tested, *pme75* identified changes in *Eco*RI, *Bam*HI and *Hind*III fragments in *T*^{Wis} DNA. *T*^{Wis}, a gain-of-function allele of *T*, arose spontaneously during a backcross of chromosome 17 from inbred strain A/J onto the inbred strain 129Sv (ref. 24). Probes F and G flank the genomic region hybridizing to clone *pme75* and detect identical fragments in DNA from strains 129Sv, A/J and a *T*^{Wis} heterozygote (data not shown). Therefore the changes in *T*^{Wis} are confined to fragments identified by the cDNA *pme75*.

To examine if these changes are responsible for the mutant phenotype of *T*^{Wis}, we constructed a genomic library from DNA of a *T*^{Wis}/+ mutant mouse in the λ vector EMBL3 (ref. 25), and isolated clones covering the *T*^{Wis} form of the *me75* gene. Restriction mapping of the *T*^{Wis} allele showed that the new

restriction fragments are caused by an insertion of 5.5 kb of foreign DNA at about half the length of the transcription unit of the *me75* gene (Fig. 3b). Sequence analysis of the 5' junction between the *me75* gene and the inserted DNA, and comparison with cDNA and genomic sequences of this gene, showed that the insertion alters a splice donor site (Fig. 3c). The splice site of the normal gene has a seven-base homology with the consensus sequence, whereas the homology decreases to five bases in the *T*^{Wis} allele. Probably the altered splice donor site is not functional, resulting in a change of the *T*^{Wis} mRNA transcript and protein. We conclude that this alteration is responsible for the mutant phenotype of *T*^{Wis}, and that *me75* is identical to the *T* gene.

The insertion would leave 352 of 436 amino-acid residues of the normal protein intact, with a modified C-terminal end due to fusion with the inserted DNA sequences. Alternatively, upstream splice donor sites would be used, shortening the protein product by one or more exons. The expected production of a modified protein is in obvious agreement with the mutant phenotype of *T*^{Wis} (*T*^{Wis}/+ animals are usually tailless). The *T*^{Wis} allele must encode a product that either competes with the wild-type product or forms an inactive complex with it, leading to an enhancement of the mutant phenotype beyond that observed for null alleles (*T*/+ animals have short tails).

One end of the inserted DNA in *T*^{Wis} has almost complete sequence homology with the U5 region of the 3' end of a murine E.Tn (for early transposon-like) element²⁶. The E.Tn sequence is a middle repetitive element of the mouse genome structurally

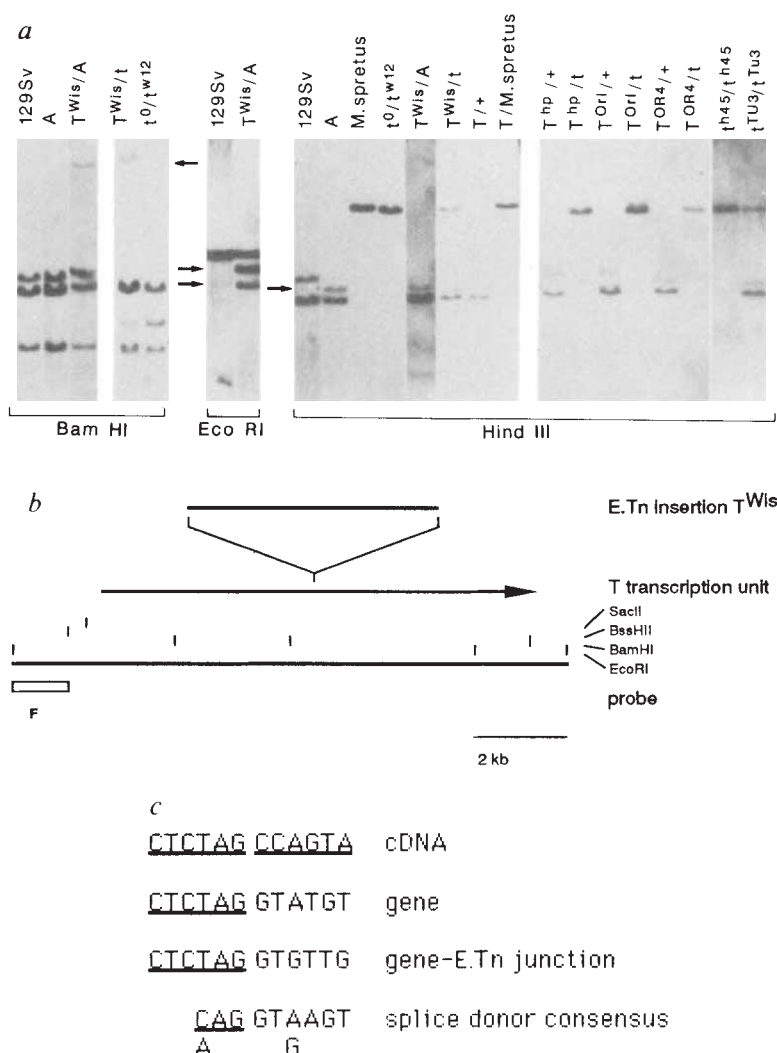


FIG. 3 Analysis of several mutant alleles. **a**, Autoradiograph of a Southern blot hybridized with the cDNA clone pme75. Arrows mark restriction fragments in the mutant chromosome T^{Wis} that differ from the control (as for BamHI and EcoRI) or mark a wild-type fragment which is changed in T^{Wis} and co-migrates with the smaller wild-type band (as for HindIII). Inbred strains 129Sv and A serve as controls because the genetic material of T^{Wis} is exclusively derived from these strains²⁴. Clone pme75 is absent in the deficiency alleles T^{hp}, T^{Orl} and T^{OR4}, and duplicated in t^{Tu3} (refs 1 and 17). In contrast to t^{Tu3}, a very similar duplication chromosome, t^{h45} (ref. 14), does not complement the T mutation. As expected, pme75 only detects one fragment (the t haplotype form) in DNA from t^{h45}/t^{h45} homozygotes, and therefore t^{h45} contains only one copy of the T gene. **b**, Restriction map of the T gene with the approximate site of insertion of a 5.5-kb E.Tn element in the T^{Wis} allele indicated on top. Probe F marks the position of the T gene on the cosmid cluster 190 (Fig. 1c). **c**, Sequence comparison between part of the T gene, the cDNA and the junction of gene and inserted DNA in T^{Wis}. A splice donor site of the T gene is changed by the insertion of the E.Tn element in such a way that the seven-base homology of the normal splice donor site with the consensus sequence decreases to a five-base homology. Exon sequences are underlined.

METHODS. Complementary DNA clones ($n=28$) were isolated out of 1.2×10^6 plaques of a library prepared from RNA of 8.5-days-post-coitum wild-type embryos²³ and screened with the cosmid 190 as probe. Repetitive sequences were removed from the probe by hybridization to total mouse DNA in solution (see Fig. 1). Southern blotting¹³, probe labelling³⁴ and hybridization³² were as previously described. The T^{Wis} gene was isolated from a genomic library constructed as described previously³⁸ in λ EMBL3 (ref. 25), and restriction mapped^{37,39} and sequenced^{40,41} as previously described.

related to retroviral sequences, and is highly transcribed in early embryos and in differentiating teratocarcinoma cells. Like retroviruses, E.Tn elements can act as mutagens²⁷.

Molecular properties of the T gene

Hybridization of the pme75 cDNA to DNA from several vertebrates at high stringency showed a high degree of conservation of the T gene sequence (Fig. 4). Drosophila DNA did not hybridize under these conditions (data not shown).

We determined the complete nucleotide sequence of clone pme75 (Fig. 5) and partial sequences of two other cDNA clones that extend less at the 5' end. We identified a single long open reading frame of 436 amino-acid residues, preceded by two in-frame stop codons at bases 13–15 and bases 49–52. The predicted T gene product does not have significant homology with any known protein stored in the PIR20 and NBRF databases. It does not seem to have a signal peptide or a potential membrane-spanning segment, but it contains six putative glycosylation sites of the canonical sequence Asn-X-Ser/Thr. It is rich in serine (13%) and proline (9%) residues.

As described in the accompanying paper by Wilkinson *et al.*¹¹, the T gene is exclusively transcribed in tissues most affected by the T mutation, early mesoderm and the notochord. The combined data indicate that the T gene encodes a novel protein involved in establishing these tissues.

Discussion

The formation of mesoderm is an important process in embryogenesis. In the mouse, mutations at several loci required in this

process have been identified. We have described here the use of a deterministic approach (from the phenotype to the gene) to isolate one of these loci, T, involving a combination of molecular and genetic techniques. This demonstrates that murine genes only known by the phenotype of a mutation are accessible to cloning.

In summary, we identified the T gene through the analysis of several T mutations. Two alleles, the original isolate T and T²¹ have small chromosomal deletions of 160–200 kb and 80–

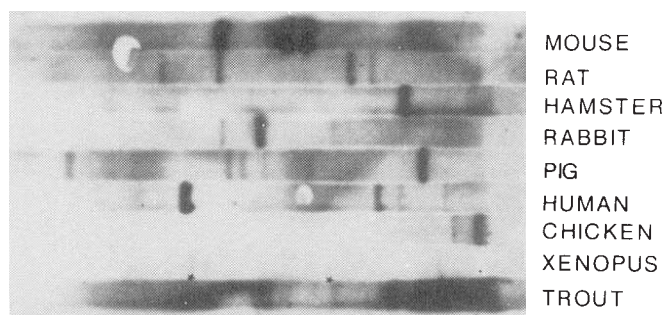


FIG. 4 Autoradiograph of a Southern blot of DNA from a variety of vertebrates hybridized with the cDNA pme75.

METHODS. DNA was cut with BamHI. Southern blotting and probe labelling were as previously described^{13,34}. Filters were hybridized and washed as described at 65 °C and 60 °C, respectively³².

FIG. 5 Nucleotide sequence and open reading frame of the cDNA *pme75* (one-letter amino-acid code). ▽, Site of insertion of an E.Tn element in the allele *T^{Wis}*. Signal sequences for potential *N*-linked glycosylation sites and polyadenylation are underlined. The 5' end of the transcript has not been identified; a few bases could be missing from the cDNA sequence shown. But the length of the cDNA equals the length of the mRNA found by northern blot analysis¹¹. Two in-frame stop codons precede the single long open reading frame starting with the first ATG in the sequence. DNA was sequenced as previously described^{40,41}.

```

GGCTCCGCAGAGTGACCCCTTTTCTTGAAAAAG
34  CGGTGGCGAGAGAAGTGAAGTGGCTGTGGGTAGGGAGTCAAGACTCCTGGAGGTGGAGGGTGGCGGGAGG
    M  S  S  P  G  T  E  S  A  G  K  S  L  Q  Y  R  V  D  H
109  ATG AGC TCG CCG GGC ACA GAG AGC GCA GGG AAG AGC CTG CAG TAC CGA GTG GAC CAC
    L  L  S  A  V  E  S  E  L  Q  A  G  S  E  K  G  D  P  T
166  CTG CTC AGC GCC GTG GAG AGC GAG CTG CAG GCG GGC AGC GAG AAG GGA GAC CCC ACC
    E  R  E  L  R  V  G  L  E  E  S  E  L  W  L  R  F  K  E
223  GAA CGC GAA CTG CGA GTG GGC CTG GAG GAG AGC GAG CTG TGG CTG CGC TTC AAG GAG
    L  T  N  E  M  I  V  T  K  N  G  R  R  M  F  P  V  L  K
280  CTA ACT AAC GAG ATG ATT GTG ACC AAG AAC GGC AGG AGG ATG TTC CCG GTG CTG AAG
    V  N  V  S  G  L  D  P  N  A  M  Y  S  F  L  L  D  F  V
337  GTA AAT GTG TCA GGC CTG GAC CCC AAT GCC ATG TAC TCT TTC TTG CTG GAC TTC GTG
    T  A  D  N  H  R  W  K  Y  V  N  G  E  W  V  P  G  G  K
394  ACG GCT GAC AAC CAC CGC TGG AAA TAT GTG AAC GGG GAG TGG GTA CCT GGG GGC AAA
    P  E  P  Q  A  P  S  C  V  Y  I  H  P  D  S  P  N  F  G
451  CCA GAG CCT CAG GCG CCC AGC TGC GTC TAC ATC CAC CCA GAC TCG CCC AAT TTT GGG
    A  H  W  M  K  A  P  V  S  F  S  K  V  K  L  T  N  K  L
508  GCC CAC TGG ATG AAG GCG CCT GTG TCT TTC AGC AAA GTC AAA CTC ACC AAC AAG CTC
    N  G  G  G  Q  I  M  L  N  S  L  H  K  Y  E  P  R  I  H
565  AAT GGA GGG GGA CAG ATG TTA AAC TCC TTG CAT AAG TAT GAA CCT GGG GAT CAC
    I  V  R  V  G  G  P  Q  R  M  I  T  S  H  C  F  P  E  T
622  ATC GTG AGA GTT GGG GGC CCG CAA CGC ATG ATC ACC AGC CAC TGC TTT CCC GAG ACC
    Q  F  I  A  V  T  A  Y  Q  N  E  E  I  T  A  L  K  I  K
679  CAG TTC ATA GCT GTG ACT GCC TAC CAG AAT GAG GAG ATT ACA GCC CTT AAA ATT AAA
    Y  N  P  F  A  K  A  T  T  L  D  A  K  E  R  N  D  H  K  D
736  TAC AAC CCA TTT GCT AAA GCC TTC CTT GAT GCC AAA GAA AGA AAC GAC CAC AAA GAT
    V  M  E  E  P  G  D  C  Q  Q  P  G  Y  S  Q  W  G  W  L
793  GTA ATG GAG GAA CCG GGS GAC TGC CAG CAG CCG GGG TAT TCC CAA TGG GGG TGG CTT
    V  P  G  A  G  T  L  C  P  P  A  S  S  H  P  Q  F  G  G
850  GTT CCT GGT GCT GGC ACC CTC TGC CCG CCT GCC AGC TCC CAC CCT CAG TTT GGA GGC
    S  L  S  L  P  S  T  H  G  C  E  R  Y  P  A  L  R  N  H
907  TCG CTC TCT CTC CCC TCC ACA CAC GGC TGT GAG AGG TCA CCA GCT CTA AGG AAC CAC
    R  S  S  P  Y  P  S  P  Y  A  H  R  N  S  S  P  T  Y  A
964  CGG TCA TCG CCC TAC CCC AGC CCC TAT GCT CAT CGG AAC AGC TCT CCA ACC TAT GCG
    D  N  S  S  A  C  L  S  M  L  Q  S  H  D  N  W  S  S  L
1021  GAC AAT TCA TCT GCT TGT CTG TCC ATG CTG CAG TCC CAT GAT AAC TGG TCT AGC CTC
    G  V  P  G  H  T  S  M  L  P  V  S  H  N  A  S  S  P  P  T
1078  GGA GTG CCT GGC CAC ACC AGC ATG CTG CCT GTG AGT CAC GCT AGC CCA CCA CTT ACT
    G  S  S  Q  Y  P  S  L  W  S  V  S  N  G  T  I  T  P  G
1135  GGC TCT AGC CAG TAT CCC AGT CTC TGG TCT GTG AGC AAT GGT ACC ATC ACC CCA GGC
    S  Q  T  A  G  V  S  N  G  L  G  A  Q  F  F  R  G  S  P
1192  TCC CAG ACA GCT GGG GTG TCC AAC GGG CTG GGA GCT CAG TTC TTT CGA GGC TCC CCT
    A  H  Y  T  P  L  T  H  T  V  S  A  A  T  S  S  S  S  G
1249  GCA CAT TAC ACA CAG GTC CAC AGC GTC TCA GCT GCC AGC TCC TCG TCT TCT AGT GGT
    S  P  M  Y  E  G  A  A  T  V  T  D  I  S  D  S  Q  Y  D
1306  TCT CCG ATG TAT GAA GGG GCT GCT ACA GTC ACA GAC ATT TCT GAC AGC CAG TAT GAC
    T  A  Q  S  L  L  I  A  S  W  T  P  V  S  P  P  S  M  END
1363  ACG GCC CAA AGC CTC CTC ATA GCC TCG TGG ACA CCT GTG TCA CCC CCA TCT ATG TGA
    ATTGAACCTTTCCTCCATGTGCTGAGACTTGTAACAACCGGTGTCAACTGGATCTTCTAGGCTCAAAGTGGCAGGC
1495  TCTTGGGACAGGGGAAAAATAAATAAATAAAGCTAGATACTAACAACCTCCATTTTCAAATAAGAGCAATAATAC
1570  ATGTCCTATAATCATGTTCTACAGCCTCTTGTGATACCTACAGTAGTGATATGTGCTACATTATGAAGCCA
1645  AGGACAGAGAGACGGCTGTGGTCCAGTTTTTTGTGACTGGCAGTTAATCAGAGTCCITTTGCTAGGTAGGGTCCCTA
1720  TATCTTGTGTTTCTCTACACATATATGTGACTTTGAAATCTGGAATTCGTCCACCCCTGTCTACTTTAGTG
1795  AGACACAAAGGTACACCTTAATGTCCCTCTGTTGCTTTAGAGTAGTTAACTTTTGGGACAGAAAAAGCATAG
1870  CCAGAAGATTGAACTGAACCGTCAACTGTTCTGCCCTTGGAACTGCTACTTTAAGCACACGTAGCTTTTGTG
1945  GTTGGGAAGTCAACTGTATGGATACTTTTCTGTTGACAAAGTAGCCAAAGCAATCTGCAGAAAGTGTCTTCTGC
2020  ATAATAAAGGCAATATATAGCACCTGGAAAAAATAAATAAATAA

```

110 kb, respectively. In the overlap of these deletions, an interval of at most 110 kb, a gene was identified which was altered by the integration of an insertion element in the gain-of-function allele *T^{Wis}*. The phenotype of *T^{Wis}* agrees well with the mutagenic event detected. It is extremely unlikely that the phenotype of *T^{Wis}*, instead of being caused by the insertion element, is produced by a change in another gene, possibly located in the same 110-kb chromosomal piece, and that the insertion detected does not result in a mutation. Although the data described here do not provide final proof, they strongly indicate that the gene isolated is the *T* gene. Additional support comes from *in situ* hybridization data obtained by Wilkinson *et*

*al.*¹¹, which show that the isolated gene is expressed in those tissues most strongly affected by the *T* mutation. We have started to produce animals transgenic for this gene to obtain direct proof for its identity with *T*.

The *T* gene is highly conserved among vertebrates, indicating that its function is also required in other vertebrates. The predicted *T* gene product is not related to any protein in the available data libraries. It lacks a signal peptide and a membrane-spanning region, indicating that it has an intracellular or possibly nuclear localization.

It has been suggested that the gain-of-function *T* allele *T^c* encodes a product that antagonizes the wild-type function²⁸. If

this is correct, then the *T* gene product should have at least two functional domains, one of which must have been retained in the *T^c* gene product. But the available molecular and genetic data do not yet allow conclusions about the molecular properties of the *T*-encoded protein. Two genetically distinct factors are known to interact with *T*. The *t^{h7}* factor can restore the normal phenotype in *T/t^{h7}* animals²⁹. It has been suggested that this property of *t^{h7}* is due to duplication of the *T* gene, so that *t^{h7}* has a double dose of the normal *T* gene function²⁹. Our data argue against this interpretation (B.G.H., unpublished results). Instead we think that the *t^{h7}* factor is independent of, but linked to *T*, and not located in the segments deleted in *T* or *T²³*. This

would mean that another locus exists that can at least partially replace *T*. More evidence for a second genetically linked locus with a phenotype similar to that of *T* comes from a recent report³⁰. Another factor, *t-int* (genetically unlinked to *T*), enhances the *T* phenotype³¹. A molecular analysis of *T* could enable these genes to be isolated and their interaction with *T* to be studied.

We have obtained molecular tools for an analysis of the properties and function of the *T* gene. An understanding of the role of *T* in the mouse should help to elucidate one of the fundamental processes of embryogenesis, the formation of mesoderm. □

Received 30 November; accepted 8 January 1989.

- Green, M. C. in *Genetic Variants and Strains of the Laboratory Mouse* 2nd edn (eds Lyon, M. F. & Searle, A. G.) (Oxford University Press, Oxford, 1989).
- Shin, H.-S. in *Human Immunogenetics* (eds Flaherty, L. & Litwin, S.) (Dekker, New York, 1988).
- Chesley, P. *J. exp. Zool.* **70**, 429–435 (1935).
- Grüneberg, H. *J. Embryol. exp. Morph.* **6**, 424–443 (1958).
- Yanagisawa, K. O., Fujimoto & Urushihara, H. *Dev. Biol.* **87**, 242–248 (1981).
- Gluacksohn-Schoenheimer, J. *Proc. natn. Acad. Sci. U.S.A.* **30**, 134–140 (1944).
- Dobrovolskaia-Zavadskaja, N. *C. R. Soc. Biol.* **97**, 114–116 (1927).
- Johnson, D. R. *Genetics* **76**, 795–805 (1974).
- Justice, M. J. & Bode, V. C. *Mouse News Lett.* **80**, 168–169 (1988).
- Searle, A. G. *Genet. Res.* **7**, 86–95 (1966).
- Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. *Nature*, **343**, 657–659 (1990).
- Röhme, D. *et al. Cell* **36**, 783–788 (1984).
- Herrmann, B. *et al. Cell* **44**, 469–476 (1986).
- Herrmann, B. G., Barlow, D. P. & Lehrach, H. *Cell* **48**, 813–825 (1987).
- Bennett, D. *et al. Genet. Res.* **26**, 95–108 (1975).
- Styrna, J. & Klein, J. *Genet. Res.* **38**, 315–325 (1981).
- Mizuno, K., Vincek, V., Figueroa, F. & Klein, J. *Immunogenetics* **30**, 112–118 (1989).
- Poustka, A. & Lehrach, H. *Trends Genet.* **2**, 174–179 (1986).
- Poustka, A., Pohl, T. M., Barlow, D. P., Frischauf, A.-M. & Lehrach, H. *Nature* **35**, 353–355 (1987).
- Poustka, A. & Lehrach, H. *Genet. Engng* **10**, 169–193 (1988).
- Bird, A. P. *Trends Genet.* **3**, 342–347 (1987).
- Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67–75 (1984).
- Fahrner, K., Hogan, B. L. M. & Flavell, R. A. *EMBO J.* **6**, 1265–1271 (1987).
- Shedlovsky, A., King, T. R. & Dove, W. F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 180–184 (1988).

- Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. J. *Molec. Biol.* **170**, 827–842 (1983).
- Kaghad, M., Maillet, L. & Brulet, P. *EMBO J.* **4**, 2911–2915 (1985).
- Shell, B., Szurek, P. & Dunnick, W. *Molec. cell. Biol.* **7**, 1364–1370 (1987).
- MacMurray, A. & Shin, H.-S. *Genetics* **120**, 545–550 (1988).
- Lyon, M. F. & Meredith, R. *Heredity* **19**, 313–325 (1964).
- Nadeau, J. H., Varnum, D. & Burkart, D. *Genetics* **122**, 895–902 (1989).
- Artzt, K., Cockingham, J. & Bennett, D. *Genetics* **116**, 601–605 (1987).
- Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995 (1984).
- Chu, G., Vollrath, D. & Davis, R. W. *Science* **234**, 1582–1585 (1986).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).
- Litt, M. & White, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6206–6210 (1985).
- Rackwitz, H. R. *et al. Gene* **40**, 259–266 (1985).
- Zehetner, G. & Lehrach, H. *Nucleic Acids Res.* **14**, 335–349 (1986).
- Frischauf, A.-M. *Meth. Enzym.* **152**, 190–199 (1987).
- Rackwitz, H. R., Zehetner, G., Frischauf, A.-M. & Lehrach, H. *Gene* **30**, 195–200 (1984).
- Chen, E. Y. & Seeburg, P. H. *DNA* **4**, 165–170 (1985).
- Labeit, S., Lehrach, H. & Goody, R. S. *Meth. Enzym.* **155**, 166–177 (1987).

ACKNOWLEDGEMENTS. B.G.H. thanks B. L. M. Hogan for laboratory space, support and encouragement, C. Carr for technical assistance, P. W. J. Rigby and U. Schwarz for support, M. Bucan, L. Stubbs, S. Myers, D. P. Barlow and F. Poirier for continued encouragement and discussions. We thank A.-M. Frischauf for cosmid libraries, G. Zehetner for computer analyses, B. L. M. Hogan for the mouse embryo cDNA library, J. Thornton and T. Gibson for discussions, M. F. Lyon, J.-L. Guenet, K. Artzt, D. Bennett, J. Klein, F. Figueroa, H. Winking, W. Dove, T. Pohl and J. Nadeau for mice and DNA samples, R. Raymond, A. Lang, S. Topps and G. Martin for maintaining mouse stocks, and D. P. Barlow, U. Schwarz, B. L. M. Hogan, P. Goodfellow, W. Dove and A. Kispert for comments on the manuscript. This work was supported by an EMBO fellowship to B.G.H.

LETTERS TO NATURE

Ozone destruction and bromine photochemistry at ground level in the Arctic spring

B. J. Finlayson-Pitts, F. E. Livingston & H. N. Berko

Department of Chemistry and Biochemistry, California State University, Fullerton, California 92634, USA

RECENT Arctic field studies in the spring at Alert, Canada (82.5° N, 62.3° W), show that ground-level ozone concentrations fell from 30–40 parts per 10⁹ (p.p.b.) to almost 0 p.p.b. in a period of time ranging from a few hours to a few days (Fig. 1)^{1,2}. Simultaneously, the concentration of brominated species collected on cellulose filters increases dramatically, with 50% or more of this bromine attributed to gaseous compounds such as HBr^{1,3}. The dramatic change in surface ozone and increase in filterable bromine seem to be due to advection of air below the inversion layer off the polar ice cap; in this air mass, surface ozone has been depleted and filterable bromine formed from photochemical reactions involving bromine compounds¹. Here we report laboratory experiments and computer kinetic modelling studies which indicate that the photochemically active bromine compound may be nitryl bromide (BrNO₂), formed in the reaction⁴ of N₂O₅ with NaBr found in sea-salt particles. BrNO₂ will photolyse rapidly at sunrise, forming ozone-destroying bromine atoms. We propose key field experiments to test this hypothesis.

Barrie *et al.*^{1,2} attributed rapid springtime changes in ground-level O₃ and filterable bromine (f-Br) concentrations to a combination of meteorological and photochemical processes. They hypothesized that photolysis of the 2–10 p.p.t. (1 p.p.t. = 1 part in 10¹²) bromoform found in the Arctic^{5–7} produced bromine. These bromine atoms then initiated a chain destruction of surface O₃ and ultimately led to the formation of HBr, consistent with peaks in f-Br throughout the Arctic at polar sunrise². Advection from the polar ice cap of these air masses, which are prevented from mixing with O₃-rich air by an inversion, results in the observed low O₃ and high f-Br levels at Alert. Their modelling studies, however, predicted a maximum of 25% loss of surface O₃ after three weeks in the absence of other brominated species such as HBr; even in the presence of 30 p.p.t. HBr, which could not be justified using known sources, only 60% loss of O₃ was predicted after three weeks. (A referee has noted that the absorption cross-sections for CHBr₃ used in ref. 1 were too large because of the presence of impurities; use of corrected cross-sections would predict even lower O₃ losses). Thus, photolysis of measured concentrations of bromoform is not consistent with the observation of air masses in which surface O₃ has been depleted by 90% or more (Fig. 1).

Laboratory studies have shown recently^{4,8} that BrNO and BrNO₂ are formed from p.p.m. (1 p.p.m. = 1 part in 10⁶) concentrations of NO₂ and N₂O₅ by reactions (1) and (2) respectively:



Nitrosyl bromide absorbs light in the actinic region with relatively large absorption cross-sections, forming bromine atoms with a quantum yield of unity^{9,10}. By analogy to ClNO_2 , BrNO_2 should also strongly absorb light in the actinic region and form Br and NO_2 with unit quantum yield¹¹.

Sturges¹² suggested that the NO_2 - NaBr reaction might serve as the source of bromine atoms, because NO_2 is present in both remote and polluted atmospheres and NaBr is a significant component of sea-salt aerosols. Our recent experiments indicate, however, that this reaction is too slow at the low NO_2 levels^{1,2,13} found in the Arctic (<50 p.p.t.). We followed the rates of formation of gaseous ClNO and BrNO in the reaction of 600 p.p.m. of NO_2 with a mixture of solid NaCl and NaBr using Fourier-transform infrared spectroscopy. In these initial studies, NaBr reacted about two orders of magnitude faster than NaCl (F. E. L. and B. J. F.-P., unpublished data). The reaction efficiency, ϕ , for the NaCl reaction has been reported¹⁴ to be only 10^{-6} – 10^{-7} , indicating that $\phi \approx 10^{-4}$ – 10^{-5} for the NO_2 - NaBr reaction.

Using this estimate of $\phi_{\text{NO}_2\text{-NaBr}}$, we included the potential concentrations of BrNO in the Arctic and its photolysis in a computer kinetic model for remote regions, incorporating bromine chemistry similar to that of Barrie *et al.*¹. The BrNO yield was estimated using gas kinetic theory to calculate¹⁵ the collision rate, R_c , between NO_2 and airborne sea-salt particles of surface area A : $R_c = CA(RT/2\pi M)^{1/2}$, where M is the molecular weight of the gas, C is its concentration, R is the gas constant and T is the absolute temperature. We calculated the amount of surface bromine available for reaction using published particle composition data³ for Alert during the winter, assuming that all of the sodium (292 ng m^{-3}) was in the form of sea salt, and that the Br/Na ratio was that of bulk seawater, 6.2×10^{-3} by mass. Based on the reported particle size distributions for Na^+ and Cl^- in the Arctic¹⁶, we assumed the sea-salt particles to be $2.5 \mu\text{m}$ in diameter. Assuming $\phi_{\text{NaBr-NO}_2} = 10^{-4}$ and a NO_2 concentration of 50 p.p.t. (ref. 13), the rate of formation of BrNO is $8 \times 10^{-3} \text{ molecule cm}^{-3} \text{ s}^{-1}$. During a typical

16-day transit time over the Arctic¹⁷, only 4×10^{-4} p.p.t. BrNO would be formed, far too low a concentration to lead to the observed surface O_3 depletion; even considering 144 days of Arctic winter, the maximum BrNO concentration is only 3×10^{-3} p.p.t.

Experiments on the N_2O_5 - NaBr reaction, combined with computer kinetic modelling, however, support the hypothesis that the photolysis of BrNO_2 formed by reaction (2) may cause the O_3 depletion. N_2O_5 is formed from the reaction of O_3 with NO_2 , which produces NO_3 . The nitrate radical, which has been measured in remote as well as polluted urban areas¹⁵, reacts with NO_2 to form N_2O_5 . From the equilibrium constant¹¹ for the NO_3 - NO_2 - N_2O_5 system and the measured levels of NO_3 and NO_2 , concentrations of N_2O_5 as high as ~ 10 p.p.b. have been calculated for the relatively moist, polluted atmosphere of southern California¹⁸. Thus, it would be surprising if some N_2O_5 were not present in the dry, remote Arctic winter.

Figure 2a shows the infrared spectrum of the gaseous products of the reaction of N_2O_5 with NaBr . Bands from unreacted N_2O_5 , BrNO , NO_2 and HNO_3 (present as an impurity in the N_2O_5), as well as the recently identified⁴ BrNO_2 , are seen. After 22 hours in the dark, BrNO_2 is gone, BrNO has decreased and NO_2 has increased (Fig. 2b).

Mass spectrometric analysis (described in ref. 19) of the products of the N_2O_5 - NaBr reaction after 22 hours in the dark at the same conditions as those of the experiment shown in Fig. 2a shows peaks at mass-to-charge ratios of 79, 81, 158, 160 and 162, attributable to bromine with an average yield in three experiments of $97 \pm 11\%$. Bromine is probably formed by the thermal decay of BrNO_2 in the gas phase, on the surface of the

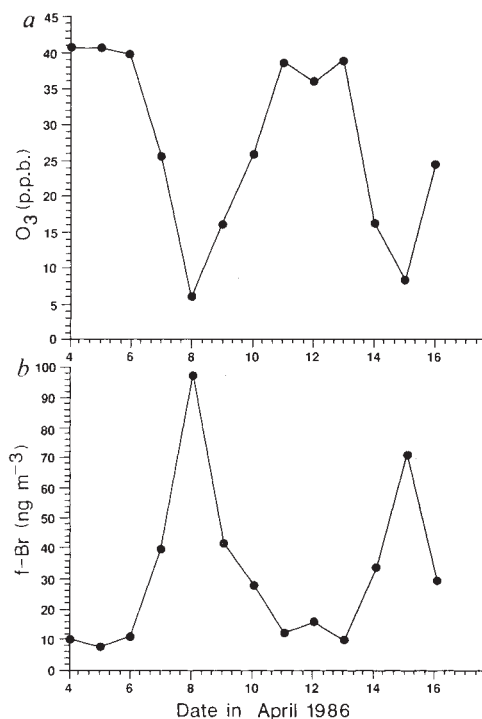


FIG. 1 Concentration-time profiles for a, O_3 and b, $f\text{-Br}$ at Alert, Canada in April 1986 as reported by Barrie *et al.*¹. These represent a combination of chemical transformations and transport of an air mass off the polar ice cap below the surface inversion.

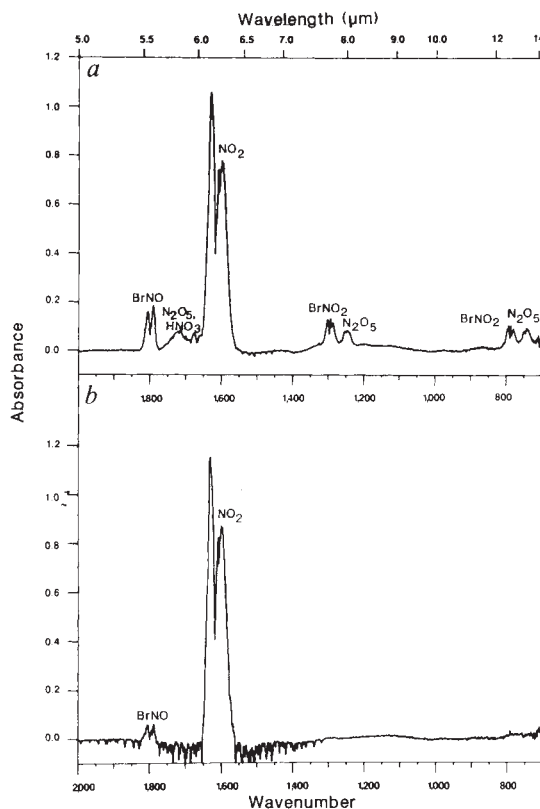


FIG. 2 Infrared spectra of the mixture of gaseous products and unreacted N_2O_5 from the reaction of 85 p.p.m. N_2O_5 with 100 g dry NaBr for a 7.5-minute reaction time at 298 K in one atmosphere of helium. Six minutes after the gas mixture was expanded into the long-pathlength cell, the spectrum in a was recorded. After 22 hours in the cell, the spectrum in b was obtained. Spectra were recorded using a multi-pass cell with White cell optics (0.8-m base path, 40-m total pathlength) and a Mattson Cygnus spectrometer at 2-cm^{-1} resolution.

reaction cell, or both, analogous to that observed^{9,10} for BrNO:



We can obtain an order-of-magnitude estimate of BrNO₂ yields based on the related reaction^{20,21} of N₂O₅ with HCl in ice in which the reaction efficiency increases²¹ from 0.03 to 0.07 as the mole fraction of HCl increases from 0 to 0.04 at 195 K. As NaBr is more reactive than NaCl, at least with NO₂, the efficiency for the N₂O₅-NaBr reaction may approach unity. Assuming a $\phi_{\text{N}_2\text{O}_5-\text{NaBr}}$ of unity and an N₂O₅ concentration of 50 p.p.t., the rate of formation of BrNO₂ for 2.5- μm sea-salt particles is 1×10^2 molecule $\text{cm}^{-3} \text{s}^{-1}$. In 144 days of Arctic winter, the theoretical yield of BrNO₂ is 43 p.p.t.; if the transit time for the air mass across the Arctic is of the order of 16 days¹⁷, ~5 p.p.t. BrNO₂ would be formed. Distributing the measured salt loading amongst smaller 1- μm particles would yield 12–108 p.p.t. of BrNO₂ in 16–144 days. If under Arctic conditions BrNO₂ decays to Br₂ as it is formed, or if the brominated products consist of a mixture of BrNO₂ and BrNO as seen in the laboratory experiments (Fig. 2a), the conclusions reached are not altered because both Br₂ and BrNO also photolyse rapidly to form bromine atoms.

The reaction of N₂O₅ with NaBr must compete with other reactions such as those with water¹⁵, ice^{20,21} and NaCl¹⁹. The continuous formation of N₂O₅ from the measured Arctic O₃ and NO₂ concentrations^{1,2,13} of ~40 p.p.b. and 25 p.p.t. respectively (rate constant $k = 6.4 \times 10^{-18} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ at 250 K; ref. 11) would yield a steady-state N₂O₅ concentration of 50 p.p.t. if the total rate of removal of N₂O₅ by all processes were 5.6×10^3 molecule $\text{cm}^{-3} \text{s}^{-1}$. As shown above, at the salt-particle concentrations measured in the Arctic³, the rate of BrNO₂ formation and hence the rate of N₂O₅ loss due to the N₂O₅-NaBr reaction for 2.5- μm particles is 1×10^2 molecule $\text{cm}^{-3} \text{s}^{-1}$. Thus, at a

steady-state N₂O₅ concentration of 50 p.p.t., only 1.8% of the total N₂O₅ loss is through the reaction with NaBr; the remaining 98% of the N₂O₅ loss could occur by other reactions such as those with water¹⁵, ice^{20,21} and NaCl¹⁹.

We used the computer kinetic model to predict the concentration-time profiles of surface ozone and filterable bromine for a solar zenith angle of 76° (typical of Alert, for example, on 6 April) and a temperature of 250 K. The model was first run for the case treated by Barrie *et al.*¹ in which photolysis of 8 p.p.t. CHBr₃ with a rate constant of $6 \times 10^{-7} \text{ s}^{-1}$ was the sole source of bromine atoms. The f-Br was calculated as the sum of HBr along with the much smaller concentrations of bromine nitrate. (Because ClONO₂ is collected on cellulose filters²², it is highly likely that BrONO₂ will be as well; M. J. Molina, personal communication). The model was then run assuming that, in addition to CHBr₃, BrNO₂ was present at 20, 50 or 100 p.p.t. respectively. As seen in Fig. 3, if bromoform is the sole bromine-atom source, the predicted O₃ destruction is slow, in agreement with the modelling results of Barrie and coworkers¹. Initial BrNO₂ concentrations in the p.p.t. range, however, give a rapid loss of surface O₃ with a corresponding increase in f-Br.

This model treats the chemistry in a closed box of air trapped below the surface inversion over the Arctic ice cap. Because of the transport of air masses, the field observations at Alert (Fig. 1) represent, in effect, a periodic sampling of this air. Thus, the photochemistry leading to the field observations may occur over a period of days to weeks before transport of the air mass to Alert, and even low p.p.t. levels of BrNO₂ could explain the observed O₃ depletion and increase in f-Br.

Even after the initial photolysis of BrNO₂, subsequent O₃ destruction is predicted because of further reactions of bromine sequestered as BrO. For example, the model predicts that two days after the photolysis of 50 p.p.t. BrNO₂, 26 p.p.t. of the bromine resides in HBr and 22 p.p.t. in BrO. At an OH concentration of 7×10^5 molecule cm^{-3} , ~75% of the HBr formed initially will react in two days to regenerate bromine atoms, and subsequently BrO radicals. Assuming that the remaining 25% of the HBr has undergone deposition and that the small amounts of atomic and molecular bromine present after two days also form BrO, a total BrO of 43 p.p.t. will be available to re-initiate O₃ loss. For example, the dotted curves shown in Fig. 3 for days 8–10 are predicted for O₃ and f-Br under these conditions. Thus, depletion of surface O₃ and generation of f-Br will continue in an air mass until all of the bromine has been removed by deposition of HBr.

The proposed identity of BrNO₂ as the cause of the rapid ozone destruction meets the necessary criteria summarized by Sturges and Barrie³: (1) the amount of filterable bromine peaks in the spring and is near zero in the summer; (2) the rapid O₃ destruction and increase in f-Br is a geographically widespread phenomenon in the Arctic; (3) f-Br increases earlier in the spring and is higher at the more southerly sampling locations at Mould Bay and Igloolik than at Alert; and (4) the bromine source is probably marine. Thus N₂O₅ will be present in significant concentrations only during the winter, because the precursor NO₃ photolyses rapidly¹⁵. The ubiquitous nature of sea salt in the Arctic^{3,16} and the higher salt concentrations at more southern locations³ indicate that the formation of BrNO₂ will occur over a large area, with concentrations being higher in more southerly regions which have higher salt loadings.

A critical field test of this hypothesis is the determination of N₂O₅ concentrations during the Arctic winter. This requires simultaneous measurements of NO₂ and NO₃; from these and the equilibrium constant¹¹ for their reaction to form N₂O₅, the concentration of N₂O₅ can be calculated¹⁵. In addition, field measurements of individual inorganic bromine compounds during the Arctic winter and spring are needed.

The reaction of BrO with HO₂ is included in both our model and that of Barrie and coworkers¹. For consistency with Barrie *et al.*¹, we assumed that it generates HOBr, which photolyses

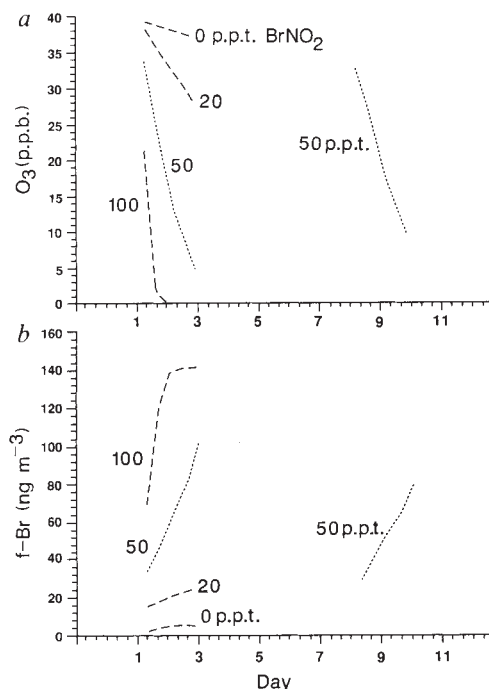


FIG. 3 Predicted concentration-time profiles for a, O₃ and b, f-Br (calculated as the sum of HBr and BrONO₂) in an air mass initially containing 8 p.p.t. CHBr₃ and 0, 20, 50 or 100 p.p.t. BrNO₂ respectively, at a solar zenith angle of 76° and temperature of 250 K. The curves for days 8–10 are the predictions of O₃ loss and formation of f-Br if 43 p.p.t. BrO formed by the initial photolysis of 50 p.p.t. BrNO₂ and subsequent reactions of the bromine atoms are available to re-initiate the ozone-destroying chain reactions (see text).

immediately. It has been recommended²³, however, that absorption cross-sections for HOCl be used for HOBr, but with the absorption spectrum shifted by 30 nm; the photolysis of HOBr would then be sufficiently slow that bromine would become sequestered as HOBr, and the chain destruction of O₃ would be much less severe. Hence experiments to measure the absorption cross-sections and quantum yields for HOBr, as well as BrNO₂, are underway in our laboratory. □

Received 28 August; accepted 22 December 1989.

- Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, P. J. & Rasmussen, R. A. *Nature* **334**, 138–141 (1988).
- Barrie, L. A., den Hartog, G., Bottenheim, J. W. & Landsberger, S. *J. Atmos. Chem.* (in the press).
- Sturges, W. T. & Barrie, L. A. *Atmos. Environ.* **22**, 1179–1194 (1988).
- Finlayson-Pitts, B. J., Livingston, F. E. & Berk, H. N. *J. Phys. Chem.* **93**, 4397–4400 (1989).
- Berg, W. W., Sperry, P. D., Rahn, K. A. & Gladney, E. S. *J. Geophys. Res.* **88**, 6719–6736 (1983).
- Berg, W. W., Heidt, L. E., Pollock, W., Sperry, P. D. & Cicerone, R. J. *Geophys. Res. Lett.* **11**, 429–432 (1984).

- Cicerone, R. J., Heidt, L. E. & Pollock, W. H. *J. Geophys. Res.* **93**, 3745–3749 (1988).
- Finlayson-Pitts, B. J. & Johnson, S. N. *Atmos. Environ.* **22**, 1107–1112 (1988).
- Houel, N. & Van Den Bergh, H. *Int. J. Chem. Kinet.* **9**, 867–874 (1977).
- Uthman, A. P. et al. *J. Phys. Chem.* **82**, 2252–2257 (1978).
- DeMore, W. B. et al. *Jet Propulsion Laboratory Publ.* No. 87–41 (September 15, 1987).
- Sturges, W. T. *Atmos. Environ.* **23**, 1167–1168 (1989).
- Bottenheim, J. W., Gallant, A. G. & Brice, K. A. *Geophys. Res. Lett.* **13**, 113–116 (1986).
- Sverdrup, G. M. & Kuhlman, M. R. *Stud. Envir. Sci.* **8**, 245–248 (1980).
- Finlayson-Pitts, B. J. & Pitts, J. N. Jr *Atmospheric Chemistry: Fundamentals and Experimental Techniques* (Wiley, New York, 1986).
- Hoff, R. M., Leitch, W. R., Fellin, P. & Barrie, L. A. *J. Geophys. Res.* **88**, 10947–10956 (1983).
- Patterson, D. E. & Husar, R. B. *Atmos. Environ.* **15**, 1479–1482 (1981).
- Atkinson, R., Winer, A. M. & Pitts, J. N. Jr *Atmos. Environ.* **20**, 331–339 (1986).
- Finlayson-Pitts, B. J., Ezell, M. J. & Pitts, J. N. Jr *Nature* **337**, 241–244 (1989).
- Tolbert, M. A., Rossi, M. J. & Golden, D. M. *Science* **240**, 1018–1021 (1988).
- Leu, M.-T. *Geophys. Res. Lett.* **15**, 851–854 (1988).
- Gandrud, B. W. et al. *J. Geophys. Res.* **94**, 11285–11297 (1989).
- Baulch, D. L. et al. *J. Phys. Chem. Ref. Data* **11**, 327–496 (1982).

ACKNOWLEDGEMENTS. We thank J. N. Pitts Jr for discussions and comments on the manuscript, R. Gill and M. Reilly for technical assistance and F. Lurmann and the Statewide Air Pollution Research Center of the University of California, Riverside for the chemical kinetics software package. We acknowledge the financial support of the NSF.

Dependence of catalytic rates on catalyst work function

C. G. Vayenas, S. Bebelis & S. Ladas

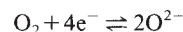
Institute of Chemical Engineering and High Temperature Chemical Processes, Department of Chemical Engineering, University of Patras, Patras GR 26110, Greece

THE catalytic activity and selectivity of metals can be altered dramatically and reversibly by supplying or removing oxide anions, O^{2−}, at the metal catalyst surface by interfacing the catalyst with an O^{2−} conductor such as zirconia^{1–6}. This effect, termed non-faradaic electrochemical modification of catalytic activity (NEMCA), leads to a steady-state catalytic reaction rate increase up to 3×10^5 times higher than the rate of supply or removal of O^{2−} at the catalyst surface^{1,3,4}. Here we report that β"-Al₂O₃, which is a Na⁺ conductor, can also induce the NEMCA effect. Furthermore we show that the origin of the NEMCA effect lies in the controlled variation of catalyst work function on polarization of the metal–solid–electrolyte interface, and demonstrate that the rates of metal-catalysed reactions depend exponentially on the average catalyst work function. Thus the influence of catalyst electronic structure, rather than surface geometry, is here the key factor in catalytic activity.

Controlled modification of the electronic and concomitant catalytic properties of metal and metal-oxide catalysts has been

a long-sought goal since the 1950s, when the importance of the electronic factor in heterogeneous catalysis was widely recognized⁷. The term 'electronic factor' indicates that the catalytic activity of metals, alloys and semiconductors is intimately related to their electronic properties. In its simplest form, this approach argues for a relationship between catalytic activity and electronic structure of the solid. A common alternative is the 'geometric factor', which attempts to relate catalytic activity to lattice spacing or, in general, surface microtopography. The catalytic properties of solid catalysts can be modified by doping the active phase or by metal–support interactions^{8–12}. It is frequently accepted that such modifications of catalyst properties result from changes in the binding energies of chemisorbed species, although geometric considerations often have an important role^{8,10}.

The solid-electrolyte cells are of the type: gaseous reactants, metal catalyst [ZrO₂(8 mol% Y₂O₃)]M, O₂. The metal M catalyses the reaction



and supplies O^{2−} to the catalyst through the gas-impervious zirconia solid electrolyte under the influence of an external voltage (Fig. 1). The increase Δ*r* in catalytic reaction rate can be up to 3×10^5 times greater than the rate $I/2F$ (where *I* is the current and *F* the Faraday constant) of supply of O^{2−} to or from the catalyst^{3–5}. We have studied the NEMCA effect on platinum, palladium and silver surfaces^{1–6}. All reactions studied so far (Table 1) exhibit the effect. We have observed reversible

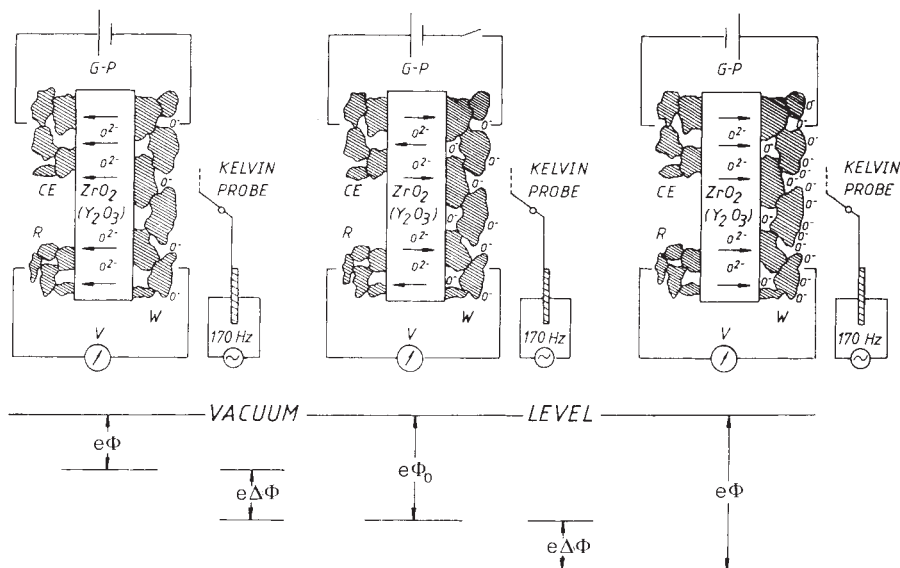


FIG. 1 Principle of NEMCA. A metal counter electrode (CE) is used to supply (right) or remove (left) oxide anions (O^{2−}) to or from the polarizable solid-electrolyte/catalyst (W) interface. The electrochemically induced spillover of oxygen anions O^{2−} produced (right) or consumed (left) at the solid-electrolyte/catalyst/gas three-phase boundaries^{1,4} causes an increase (right) or decrease (left) in the catalyst work function $e\Delta\Phi$. $e\Delta\Phi = e\Delta V_{WR}$, where ΔV_{WR} is the overpotential measured between the catalyst and the metal reference electrode (R).

TABLE 1 Catalytic reactions exhibiting the NEMCA effect

Positive (electrophobic) NEMCA effect ($\Delta r > 0$ with $I > 0^*$, $e\Delta\Phi > 0$)							
Reactants	Products	Catalyst	Electrolyte	Temperature (°C)	$\Delta r/(I/2F)^\ddagger$	r/r_0	References
C_2H_4, O_2	C_2H_4O, CO_2	Ag	$ZrO_2-Y_2O_3$	320–470	0, 300	<3	3†
C_3H_6, O_2	C_3H_6O, CO_2	Ag	$ZrO_2-Y_2O_3$	320–420	0, 420	<2	3†
C_2H_4, O_2	CO_2	Pt	$ZrO_2-Y_2O_3$	260–450	0.3×10^5	<55	1, 3, 4
C_2H_4, O_2	CO_2	Pt	$\beta''-Al_2O_3$	180–300	0.5×10^4	<4	This work
CO, O_2	CO_2	Pt	$ZrO_2-Y_2O_3$	300–550	0, 500	<3	2, 3
CO, O_2	CO_2	Pd	$ZrO_2-Y_2O_3$	400–550	0.10^3	<1.5	3
CH_3OH, O_2	H_2CO, CO_2	Pt	$ZrO_2-Y_2O_3$	300–500	0.10^4	<3	This work†, 3
CH_4, O_2	CO, CO_2, C_2H_4	Ag	$ZrO_2-Y_2O_3$	650–750	0, 5	<30	This work†
Negative (electrophilic) NEMCA effect ($\Delta r > 0$ with $I < 0$, $e\Delta\Phi < 0$)							
CO, O_2	CO_2	Pt	$ZrO_2-Y_2O_3$	300–550	0, –500	<6	2, 3
CH_3OH, O_2	H_2CO, CO_2	Pt	$ZrO_2-Y_2O_3$	300–550	0, -10^4	<15	This work†, 3
CH_3OH	H_2CO, CO, CH_4	Ag	$ZrO_2-Y_2O_3$	550–750	0, –25	<6	5†
CH_3OH	H_2CO, CO, CH_4	Pt	$ZrO_2-Y_2O_3$	400–500	0, –10	<3	3†

* Current is defined positive when O^{2-} are supplied to or Na^+ removed from the catalyst surface.

† Change in product selectivity observed.

‡ Upper and lower limits of range in measured values.

catalytic rate enhancements of up to a factor of 60 at constant coverages of adsorbed species^{3,4} with significant concomitant changes in activation energies^{4,5} and product selectivity^{3,5}. No catalyst or electrolyte deterioration problems were encountered. The catalytic rates depend exponentially on the ohmic-drop-free catalyst (working electrode: W) potential V_{WR} with respect to a reference electrode (R), and we proposed that the NEMCA effect is due to changes in the catalyst work function on polarization of the catalyst–zirconia interface.

Here we report experimental results showing that: (1) the effect is not restricted to the use of O^{2-} -conducting solid electrolytes but can also be induced using Na^+ -conducting solid electrolytes, such as $\beta''-Al_2O_3$; (2) the change ΔV_{WR} in the potential of catalysts that function as electrodes in solid-electrolyte cells is $\Delta\Phi$, where $e\Delta\Phi$ is the change in the work function of the gas-exposed catalyst surface; (3) over wide ranges of work function, typically 0.3–1 eV, the rates of catalytic reactions depend exponentially on the catalyst work function,

according to:

$$\ln(r/r_0) = \alpha e(\Phi - \Phi^*)/k_B T$$

where r_0 is the regular (that is, open-circuit) catalytic rate and α and Φ^* are reaction- and catalyst-specific constants. The constant α typically takes values between 1 and –1. Over the same range of work function, catalytic activation energies change significantly.

These observations demonstrate the validity of our semiquantitative explanation of the NEMCA effect^{1,3–5}, which is based on a uniform change of the catalyst work function induced by an electrochemically controlled spillover of ions onto the catalyst

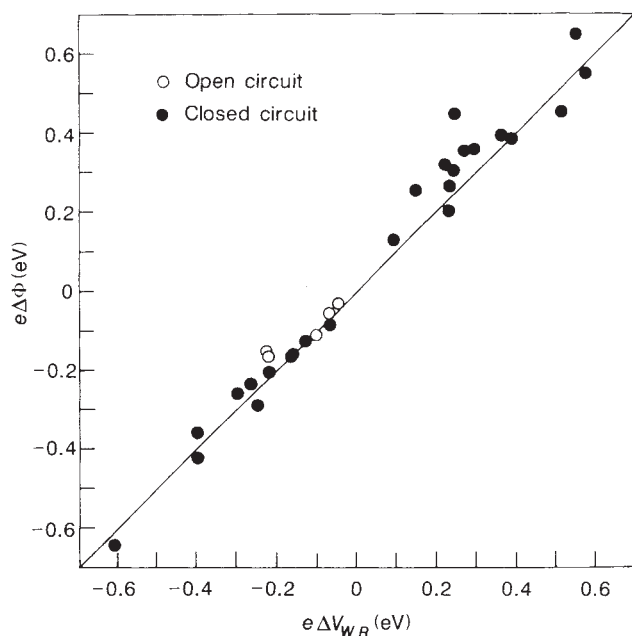


FIG. 2 Effect of change in ohmic-drop-free catalyst potential ΔV_{WR} on the work function of the gas-exposed catalyst surface; Pt film; $ZrO_2(Y_2O_3)$ solid electrolyte; $T = 300^\circ C$. The straight line of slope unity is that predicted using the theory in refs 1–5.

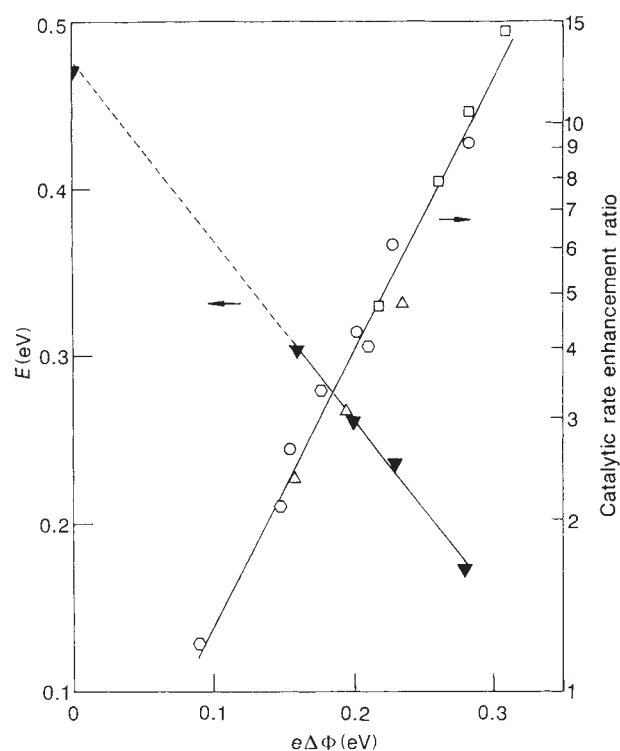


FIG. 3 Effect of catalyst work function at constant coverage of adsorbed species⁴ on the activation energy E and catalytic rate enhancement ratio r/r_0 of C_2H_4 oxidation on Pt; $P_{C_2H_4} = 0.4$ kPa, $P_{O_2} = 4.8$ kPa. $\circ$, $T = 399^\circ C$, $r_0 = 2.3$ s⁻¹; Δ , $T = 360^\circ C$, $r_0 = 1.7$ s⁻¹; $\circ$, $T = 324^\circ C$, $r_0 = 1.0$ s⁻¹; $\square$, $T = 296^\circ C$, $r_0 = 0.6$ s⁻¹. The slope of the line for the activation energy is –1 and the slope of the line for the catalytic rate is $0.6/k_B T$.

surface and on the concomitant controlled change in the strength of chemisorptive bonds^{4,5,7}. They also reveal a surprisingly simple relationship between the catalytic activity and the catalyst work function, measured *in situ* and, for the first time, varied at constant reactant composition. This relationship quantifies the importance of the electronic factor in heterogeneous catalysis.

The continuous-flow, atmospheric-pressure experimental apparatus using on-line gas chromatography, mass spectrometry and infrared spectroscopy for reactant and product analysis has been described in previous papers¹⁻⁵ where catalyst-film preparation and characterization details are also presented. The porous catalyst films that we use have thickness of $\sim 5\ \mu\text{m}$, surface areas of $100\text{--}1,500\ \text{cm}^2$ and are supported on a solid electrolyte of area $2\ \text{cm}^2$. The three-electrode system (Fig. 1) used with the current-interruption technique permits accurate measurement of the ohmic-drop-free catalyst potential V_{WR} relative to a standard reference air electrode, and of the catalyst overpotential ΔV_{WR} (refs 4, 5) which has^{1,5} a key role in the NEMCA effect. The catalyst work function was measured using a Kelvin probe (Besocke/Delta-Phi-Elektronik) with a 2.5-mm-diameter gold-grid vibrating electrode placed $\sim 500\ \mu\text{m}$ from the catalyst surface (Fig. 1). We used two types of solid electrolytes—fully stabilized zirconia (8 mol% Y_2O_3 in ZrO_2), an O^{2-} conductor, and $\beta''\text{-Al}_2\text{O}_3$, a Na^+ conductor.

Figure 2 shows that the change in the work function of the gas-exposed (that is, catalytically active) surface of the catalyst electrode is $e\Delta V_{\text{WR}}$, both under closed- and open-circuit conditions. In the former case V_{WR} was varied by changing the polarizing current with the catalyst exposed to air, whereas in the latter the catalyst was exposed to $\text{NH}_3/\text{O}_2/\text{He}$ and $\text{CO}/\text{O}_2/\text{He}$ mixtures. This result has been predicted theoretically^{1,3-5} and stems from the spatial uniformity of the Fermi level throughout the conductive catalyst film, including the catalyst-gas and the catalyst-electrolyte interfaces. Thus the e.m.f. of solid-electrolyte cells with electrodes made of the same material provides a direct measure of the difference in work function between the working (catalyst) and reference gas-exposed electrode surfaces.

Figure 3 shows the effect of catalyst work function change on the rate of the platinum-catalysed ethylene oxidation to carbon dioxide and on the corresponding activation energy. Above a threshold work function value $e\Phi^*$, the catalytic rate increases exponentially with $e\Phi$. We have also observed threshold $e\Phi^*$ values in some other studies^{1,4,5}. Their physical origin is not clear and could be related to the form of the distribution of the density of states around the Fermi level. Lateral interactions have been invoked to explain similar abrupt variations of the heat of adsorption with $e\Phi$ in ultra-high-vacuum studies¹³. Increasing $e\Phi$ weakens the platinum-chemisorbed oxygen bond, the cleavage of which is rate limiting⁴, and causes a decrease in activation energy comparable to $e\Delta\Phi$. We note that the observed $e\Delta\Phi$ changes, which are due to spillover (spreading) of oxygen anions (very probably O^- ; refs 1, 4, 14) at the catalyst surface, require only a small coverage of anions ($<5\%$ of a monolayer, as calculated using the Helmholtz equation with an ion radius of $0.17\ \text{nm}$) which does not affect significantly the coverages of covalently bonded species⁴. When $\beta''\text{-Al}_2\text{O}_3$ is used as the solid electrolyte a qualitatively similar behaviour is observed, that is, the rate of carbon dioxide formation decreases exponentially with the decrease in the work function resulting from the spillover of Na^+ onto the catalyst surface. Thus at 290°C , and oxygen and ethylene partial pressures of $5\ \text{kPa}$ and $0.02\ \text{kPa}$ respectively, a 70% reversible decrease in reaction rate is observed when the catalyst potential is decreased by $400\ \text{mV}$. This rate decrease is 48,000 times higher than the rate of supply of Na^+ .

The catalytic reactions studied so far (Table 1) fall into two groups depending on whether their rate increases or decreases exponentially with increasing $e\Phi$. In the former case of 'elec-

trophobic' reactions⁵ the rate-limiting step of the catalytic process involves cleavage of a metal/electron-acceptor adsorbate bond. In the latter case of 'electrophilic' reactions⁵ cleavage of an intra-adsorbate bond of an electron-acceptor adsorbate is rate limiting. The observed dependence of catalytic rates (exponential) and activation energies (linear) on catalyst work function provides strong evidence that, as previously proposed¹⁻⁵, NEMCA is a long-range electronic effect. Long-range effects have been predicted theoretically for some adsorbates on metal surfaces¹⁵ and have been invoked to explain the dependence of the sticking coefficient of oxygen on alkali-doped germanium, which was found to depend exponentially on the alkali-induced $e\Phi$ change, regardless of alkali type¹⁶. 'Short-range' explanations cannot yet, however, be entirely ruled out. The use of temperature-programmed desorption¹⁷ and of scanning tunnelling microscopy¹⁸ as an atomic-scale $e\Phi$ probe could further elucidate this point.

These results show that the NEMCA effect¹⁻⁶ is not limited to a particular metal catalyst or solid electrolyte and underline the importance of the electronic factor in heterogeneous catalysis by directly establishing the relationship between catalyst work function and catalytic activity. This exponential relationship seems to apply to all metal-catalysed reactions. Metal crystallites on supports may behave similarly and scanning tunneling microscopy¹⁸ could be useful here also. Our results also show that solid-electrolyte cells can be used to conveniently control catalyst work function and to influence catalytic activity and selectivity in desirable directions. $\square$

Received 11 October; accepted 29 December 1989

1. Vayenas, C. G., Bebelis, S. & Neophytides, S. *J. phys. Chem.* **92**, 5083–5085 (1988).
2. Yentekakis, I. V. & Vayenas, C. G. *J. Catal.* **111**, 170–188 (1988).
3. Vayenas, C. G., Bebelis, S., Neophytides, S. & Yentekakis, I. V. *Appl. Phys.* **A49**, 95–103 (1989).
4. Bebelis, S. & Vayenas, C. G. *J. Catal.* **118**, 125–146 (1989).
5. Neophytides, S. & Vayenas, C. G. *J. Catal.* **118**, 147–163 (1989).
6. Lintz, H.-G. & Vayenas, C. G. *Angew. Chem.* **101**, 725–732 (1989); *Int. Engl. Edn* **28**, 708–715 (1989).
7. Boudart, M. *J. Am. chem. Soc.* **74**, 3556–3561 (1952).
8. Sachtler, J. W. A. & Somorjai, G. A. *J. Catal.* **81**, 77–94 (1983).
9. Tan, S. A., Grant, R. B. & Lambert, R. M. *J. Catal.* **106**, 54–64 (1987).
10. Haller, G. L. & Resasco, D. E. *Adv. Catal.* **36**, 173–235 (1989).
11. *Catalyst Design: Progress and Perspectives* (eds Hegedus, L. L. et al.) (Wiley, New York, 1987).
12. *Studies in Surface Science and Catalysis* Vol. 11 (eds Imelik, B. et al.) (Elsevier, Amsterdam, 1982).
13. Madden, H., Küppers, J. & Ertl, G. *J. chem. Phys.* **58**, 3401–3410 (1973).
14. Arakawa, T., Saito, A. & Shiohara, J. *Appl. Surf. Sci.* **16**, 365–373 (1983).
15. Feibelman, P. J. & Hamann, D. R. *Surf. Sci.* **149**, 48–66 (1985).
16. Surnev, L. *Surf. Sci.* **11**, 458–470 (1981).
17. Bonzel, H. P. *Surf. Sci. Rep.* **8**, 43–125 (1987).
18. Binnig, G. & Rohrer, H. *Surf. Sci.* **126**, 236–244 (1983).

ACKNOWLEDGEMENTS. This work was supported by the Volkswagen Foundation (FRG) and by the EEC Non-nuclear Energy Programme.

Estimates of Antarctic precipitation

David H. Bromwich

Byrd Polar Research Center, Ohio State University, Columbus, Ohio 43210, USA

PRECIPITATION fluctuations over Antarctica are a potentially important contributor to variations in global sea level^{1,2}. Direct measurement of precipitation is, however, fraught with practical difficulties³. Two methods may be used to calculate indirectly the net flux of water (precipitation minus sublimation rate) to the surface of Antarctica: the first uses values of poleward atmospheric moisture transport obtained from climatological studies, and the second uses glaciological measurements of the accumulation rate. Here I show that the two estimates so derived are in marked disagreement for the entire continent but concur for the interior area between 80°S and the pole. I conclude that the large discrepancy near the coast is due to a calculated poleward moisture transport that is smaller than the actual value, as a result of

TABLE 1 Annual zonal averages of meridional moisture transport and its convergence

Transport ($\text{kg m}^{-1} \text{s}^{-1}$)			Convergence ($\text{kg m}^{-2} \text{yr}^{-1}$) [$\bar{P} - \bar{E}$]			
Latitude °S	PO	HR	Latitude Belt °S	PO	HR	BR
50	-17	-10.5	50-60	290	160	452
60	-10	-6.7	60-70	270	133	322
70	-3	-3.7	70-80	90	97	183
80	-1	-1.6	80-90	60	77	64
90	0	0	70-90	82	92	153

Negative meridional transports are directed poleward. Square brackets denote a zonal average. $1 \text{ kg m}^{-2} \text{yr}^{-1}$ is equivalent to 1 mm yr^{-1} . PO, 10-year atmospheric averages from Peixoto and Oort⁴; HR, 9-year atmospheric averages from Howarth⁵ and Howarth and Rayner⁶; BR, somewhat outdated climatological surface-based averages from Baumgartner and Reichel⁸.

deficiencies in evaluating the effects of cyclones and surface winds at the coast. Improvements in the climatological atmospheric database should therefore make possible reliable estimates of Antarctic precipitation variations.

The budget equations used in the calculations based on atmospheric water balance and on glaciological accumulation measurements are, respectively³,

$$\langle \bar{P} - \bar{E} \rangle = -\langle \nabla \cdot \mathbf{Q} \rangle - \left\langle \frac{\partial \bar{W}}{\partial t} \right\rangle \quad (1)$$

$$\langle \bar{P} - \bar{E} \rangle = \langle \bar{B} \rangle + \text{drift snow loss} + \text{run-off} \quad (2)$$

The angle brackets represent a spatial average, the bar denotes a time average, P precipitation, E evaporation/sublimation, B the accumulation rate and $\bar{W}$ the precipitable water. $\mathbf{Q}$ is the horizontal moisture-transport vector and is defined as

$$\mathbf{Q} = \int_0^{p_s} (q\mathbf{V})/g \, dp$$

where q is specific humidity, $\mathbf{V}$ the horizontal wind vector, g

the acceleration due to gravity and p the pressure. The integration is carried over the depth of the atmosphere from the surface ($p = p_s$) to the top ($p = 0$). W is given by $\int_0^{p_s} (q/g) \, dp$. Because climatological conditions are of interest, results are averaged over a decade or more. For annual averages the second term on the right-hand side of equation (1) is negligible⁴.

Multiannual estimates of the poleward atmospheric moisture transport in high southern latitudes (which allow the first term on the right side of equation (1) to be evaluated) are available from the global analysis of Peixoto and Oort⁴ (1963-1973) and the composited hemispheric analyses of Howarth⁵ (1973-1978) and Howarth and Rayner⁶ (1980-1984)—see Fig. 1. The data in ref. 4 were based on radiosonde observations and the data in refs 5 and 6 were based on numerical syntheses by the Australian Bureau of Meteorology of all available observations, such as data from radiosondes, aircraft and satellites. Each data set contains systematic errors and has particular strengths and weaknesses⁷.

Table 1 compares the north-south (meridional) components of the moisture transports given in refs 4-6. Out over the Southern Ocean (50° - 72° S) where the radiosonde network is extremely sparse, the transports differ substantially. Close to and over Antarctica ($\sim 72^\circ$ - 90° S), the two sets of values are very similar. The $\bar{P} - \bar{E}$ values calculated from equation (1) also illustrate the large atmospheric discrepancies over the Southern Ocean and much better agreement over Antarctica. There are large differences between the atmospheric $\bar{P} - \bar{E}$ estimates and the obsolete surface-based values of ref. 8, except for the 80° - 90° S latitudinal belt.

As will be discussed below, the accumulation contribution is by far the largest term on the right-hand side of equation (2). The recent comprehensive accumulation analysis of Giovinetto and Bentley⁹ has been chosen for this term. By careful selection among existing data sets¹⁰, they obtained slightly smaller rates over 90% of the continent (interior parts) and larger rates over the coastal margins. Overall their values for Antarctica are $\sim 8\%$ smaller than most previous estimates. Because of the emphasis on mutually consistent observations and data obtained with improved measurement techniques, I regard this analysis as the best long-term-average description of the annual accumulation rate so far.

To complete the surface-based estimate of $\bar{P} - \bar{E}$ from equation (2) the drift snow transported across the coastline and the liquid run-off into the ocean must be evaluated. Loewe¹¹ has done a careful analysis of drift snow transport over both Greenland and Antarctica, and found that it becomes a progressively smaller component of the ice-sheet mass budget as the area increases. The liquid run-off is small, and I have used the estimates in ref. 12.

In Table 2 I compare the annual average $\bar{P} - \bar{E}$ values derived from surface-based and climatological atmospheric observations for the entire continent and for the latitudinal belt 80° - 90° S. The drift snow loss and run-off for Antarctica were obtained as described above. The values for 80° - 90° S were calculated as follows. Because this area is entirely contained within the continent (Fig. 1) and spans regions with generally low accumulation rates (almost everywhere $< 15 \text{ mm yr}^{-1}$), the drift snow loss should be small and is approximated by zero. As the summer air temperatures are almost everywhere continuously below the freezing point, the run-off must be zero¹³. Table 2 shows for the entire continent that the atmospheric estimates of $\bar{P} - \bar{E}$ are only one-third to one-half the size of the surface-based estimation. The somewhat dated⁸ (surface) $\bar{P} - \bar{E}$ value for Antarctica is 141 mm yr^{-1} , slightly smaller than the surface-based estimate in Table 2. For 80° - 90° S the atmospheric and terrestrial approaches give similar results. Because the accumulation values are well determined and larger estimates of drift snow loss and run-off would only add to the huge discrepancy between surface-based and atmospheric estimates of $\bar{P} - \bar{E}$ ($155 - (80 \text{ or } 50) = 75$ or 105 mm yr^{-1}), the poleward atmospheric moisture transport

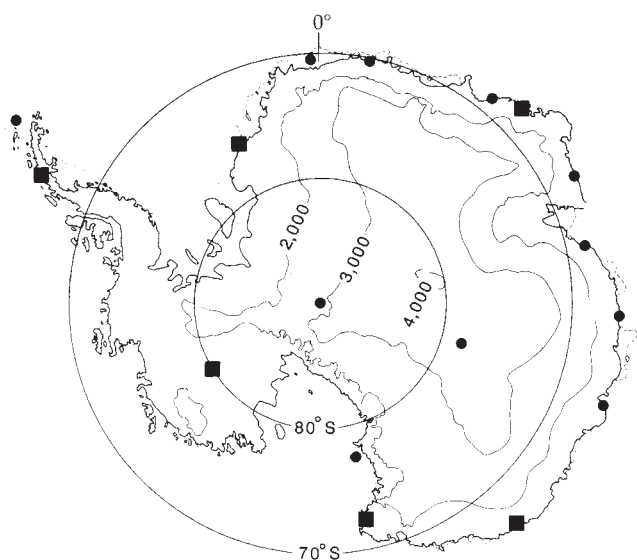


FIG. 1 Antarctic radiosonde stations, the data from which primarily determined the high-latitude atmospheric moisture fluxes obtained in ref. 4 and in refs 5, 6. Filled circles denote sites contributing to both studies and filled squares to only one. Notice the measurement gap along the West Antarctic coast (to the left); no upper-air observations have been collected at Byrd Station (80° S) since the early 1970s. Thin continuous lines are elevation contours in metres, starting at 2,000 m.

TABLE 2 Surface-based and atmospheric estimates of $\bar{P} - \bar{E}$ (mm yr⁻¹)

Area	$\langle \bar{E} \rangle$	Drift snow loss	Surface run-off	$\langle \bar{P} - \bar{E} \rangle$ from accumulation data	$\langle \bar{P} - \bar{E} \rangle$ from atmospheric data of PO, HR
Antarctica	143*	8†	0-5‡	151-156	44, 49§ 76, 85
80°-90°S	82	~0	0	82	60, 77

PO, Peixoto and Oort, ref. 4; HR, Howarth and Rayner, refs 5, 6; BR, Baumgartner and Reichel, ref. 8.

* ref. 9.

† ref. 11.

‡ ref. 12.

§ Method (1): Value for PO taken directly from ref. 22. Value for HR calculated from PO value by multiplying by ratio of HR meridional moisture transport convergence for 70°-90°S and PO value (from Table 1).

|| Method (2): Calculated from $\bar{P} - \bar{E}$ for 70°-90°S in Table 1 by multiplying by BR ratio of $\bar{P} - \bar{E}$ for Antarctica and $\bar{P} - \bar{E}$ for 70°-90°S (141 mm yr⁻¹/153 mm yr⁻¹ = 0.922).

at the coast must be responsible for the difference.

Possible factors are systematically low water-vapour values and grossly underestimated transports. Flight tests¹⁴ of two types of radiosonde humidity elements showed that relative humidity measurements during night-time were systematically lower than the actual values by fractional amounts of 0.10 to 0.15. With most Antarctic precipitation falling during and around the polar night³, this systematic error is approximately representative for Antarctica, and cannot account for the discrepancy.

There are two studies that demonstrate that atmospheric and surface-based estimates of $\bar{P} - \bar{E}$ can be in close agreement. First, Masuda¹⁵ calculated the moisture transported across 70°S from 1979 atmospheric analyses by the European Centre for Medium-Range Weather Forecasts; for that year the atmospheric circulation over the Southern Ocean was monitored substantially better than either before or, to a lesser extent, since¹⁶. His 1979 $\bar{P} - \bar{E}$ value for Antarctica (using method (2) in Table 2) was 136 mm yr⁻¹, only 15-20 mm yr⁻¹ less than the new long-term-average surface-based estimate. Second, I have derived³ the 1972 atmospheric water balance for a part of East Antarctica chosen so that the available radiosonde observations accurately monitored the poleward moisture transport across the coast. In Table 3 the atmospheric values are compared with the modern surface synthesis in ref. 9. Even though interannual variability is not included, the two sets of estimates are seen to agree within the limits of error.

The probable causes of the anomalously small calculated poleward moisture transports at the Antarctic coast can be identified. As noted by Masuda¹⁵, zonal averages of the poleward transport can be broken down into transport contributions from the transient eddies, the stationary eddies and the mean north-south air motions; it is the convergence of the poleward transport that yields $\bar{P} - \bar{E}$ values in equation (1). He found that the apparently reliable 1979 poleward transport across 70°S (dis-

cussed above) was dominated by the transient eddies (92% of the total), with the other two components nearly cancelling (+21% and -13%, respectively). Masuda compared transport components for the climatological analysis of Nakamura and Oort¹⁷ (closely related to that in ref. 4) with those for 1979. The biggest differences were that both the transient eddy and mean meridional transports increased in his study by about the same amount, such that $\bar{P} - \bar{E}$ for 70°-90°S was four times larger.

Masuda's results are consistent with the present work. First, underestimating transient-eddy transports leads to an underestimate of the poleward transport at the coast but would have little impact in the interior, because the transient-eddy transport is a large fraction of the total at the coast but decreases to small values at high elevations^{3,18}. Second, the characteristics of the dominant low-level part of the mean north-south air motions, namely the surface-wind regime¹⁹, lead me to suspect that the inability to diagnose correctly this shallow and spatially variable feature²⁰ is important for mean transport determination at the coast²¹ but negligible in the interior. Detailed exploration of this topic is, however, beyond the scope of this report.

Although recent climatological atmospheric analyses cannot be used to infer precipitation changes over Antarctica, and thus determine their impact on global sea level, results from studies for restricted areas and limited times indicate that improvements to the database and analysis methodology¹⁵ would make this monitoring possible. An effective enhancement would be to eliminate the upper-air data gap along the coast of West Antarctica. □

TABLE 3 East Antarctic $\bar{P} - \bar{E}$ values

Source	$\langle \bar{P} - \bar{E} \rangle$ (mm yr ⁻¹)		
	0°-110°E 68.4°-78.2°S	0°-55°E 69.3°-76.8°S	55°-110°E 67.2°-79.8°S
Moisture budget, ref. 3 (value ±1 standard error)	105 ± 23	131 ± 43	105 ± 23
ref. 9	108	109	108
ref. 23	144	133	151

Moisture budget, 1972 atmospheric water balance. The data in refs 9 and 23 are multiannual accumulation analyses. The values inferred from the outdated accumulation analysis in ref. 23 are presented for comparison with previous work in ref. 3. Results are given for the entire domain and two subsections.

Received 25 September; accepted 29 December 1989.

- Oerlemans, J. *Nature* **290**, 770-772 (1981).
- Zwally, H. J. *Science* **246**, 1589-1591 (1989).
- Bromwich, D. H. *Rev. Geophys.* **26**, 149-168 (1988).
- Peixoto, J. P. & Oort, A. H. in *Variations in the Global Water Budget* (eds Street-Perrott, A., Beran, M. & Ratcliffe, R.) 5-65 (Reidel, Dordrecht, 1983).
- Howarth, D. A. *Mon. Weath. Rev.* **111**, 1259-1272 (1983).
- Howarth, D. A. & Rayner, J. N. in *2nd Int. Conf. Southern Hemisphere Met.* 163-166 (Am. Met. Soc., Boston, 1986).
- Karoly, D. J. & Oort, A. H. *Mon. Weath. Rev.* **115**, 2033-2059 (1987).
- Baumgartner, A. & Reichel, E. *The World Water Balance* (Elsevier, New York, 1975).
- Giovinetto, M. B. & Bentley, C. R. *Antarct. J. U.S.* **20**(4), 6-13 (1985).
- Giovinetto, M. B., Bentley, C. R. & Bull, C. *Antarct. J. U.S.* **24**(1), 7-13 (1989).
- Loewe, F. in *Met. Dept. Publ.* 13 (University of Melbourne, 1970).
- Rubin, M. J. in *Antarctic Research* (eds Wexler, H., Rubin, M. J. & Caskey, J. E. Jr) 149-159 (Am. Geophys. Un., Washington, DC, 1962).
- Paterson, W. S. B. in *Environment of West Antarctica: Potential CO₂-Induced Changes* 140-154 (Natl. Acad. Press, Washington, DC, 1984).
- Morrissey, J. F. & Brousaides, F. J. *J. appl. Met.* **9**, 805-808 (1970).
- Masuda, K. in *Proc. 11th NIPR Symp. Polar Met. Glaciol.* (Natl. Inst. Polar Res., Tokyo, in the press).
- Hollingsworth, A. *Weather* **44**, 278-285 (1989).
- Nakamura, N. & Oort, A. H. *J. geophys. Res.* **93**, 9510-9524 (1988).
- Lettau, B. *Tellus* **21**, 331-340 (1969).
- Parish, T. R. *Rev. Geophys.* **26**, 169-180 (1988).
- Parish, T. R. & Bromwich, D. H. *Nature* **328**, 51-54 (1987).
- Bromwich, D. H. *Antarct. J. U.S.* **13**(4), 196-197 (1978).
- Bryan, F. & Oort, A. H. *J. geophys. Res.* **89**, 11717-11730 (1984).
- Bull, C. in *Research in the Antarctic. Publ.* 93 (Am. Ass. Advancement Sci., Washington, DC, 1971).

ACKNOWLEDGEMENTS. I thank M. Giovinetto and J. Zwally for their comments.

A record of Holocene summer climate from a Canadian high-Arctic ice core

R. M. Koerner* & D. A. Fisher†

* Terrain Sciences Division, Geological Survey of Canada 601 Booth Street, Ottawa, Canada, K1A 0E8

† Department of Glaciology, University of Copenhagen, Haraldsgade 6, DK-2200, Copenhagen, Denmark

MELT layers in ice cores are formed by melting on the snow-pack surface, which produces water that percolates down and refreezes in the colder snow layers below. Melt layers are distinguished by their 'bubble' texture. Because more ice forms in the snow pack in warm summers than in cool summers, the changing concentration with depth of melt layers in snow pits^{1,2} and ice cores³⁻⁵ is an indicator of past summer climate. Here we report on studies of melt layers in two Canadian high-Arctic surface-to-bedrock ice cores. Although changes in elevation of the drill site owing to isostatic depression and uplift may contribute up to 40% of the total climatic signal recorded in the melt layers, we interpret their varying concentrations in terms of changing summer temperatures over the past 10,000 years. The warmest summers occurred 8–10 kyr ago and the coldest only 150 years ago. The summers over the past 100 years have been the warmest for more than 1,000 years, but are still not as warm as those of the early Holocene. The melt-layer record is in broad agreement with geological records of glacier retreat during the Holocene⁶⁻⁸, indicating that these data also contain information of regional significance.

A previous study of the Devon Ice Cap (ref. 3, in which the methodology is described) showed that the melt record at the top of the ice cap bears strong relationships to melting over the entire ice cap, to melting on other ice caps and also to late-summer sea-ice conditions in the Queen Elizabeth Islands area. Ice-core melt-layer records are therefore not just site specific.

The two 127-m-long surface-to-bedrock cores we discuss (A-84, A-87) are from sites that are 50 m apart and located at the top of the flow line on Agassiz Ice Cap, Ellesmere Island (80.7° N, 73.1° W, 1,730-m elevation). We present here only the record from A-84 (Fig. 1a) because the A-87 core has sections in it that were affected by malfunctions of the thermal drill used to obtain the cores. The overall records from both cores are the same for the resolution used here. A detailed analysis of the Holocene and pre-Holocene stable-isotope records for two other cores drilled 1 and 2 km downslope has already been published⁹⁻¹².

We express the data in the form of melt-ice percentage (P) and any firm in the core is corrected to its ice equivalent. The core densities were measured at the drill site. P is dependent not only on the amount of melting each summer but also on the annual accumulation rate. Thus, for example, if the accumulation rate increased but the amount of melting remained the same, the value of P would decrease. Theoretical considerations¹³ as well as the ice-core records in both Antarctica¹⁴ and Greenland^{15,16} indicate, however, that accumulation rates usually increase with warmer temperatures and decrease with cooler temperatures. Therefore warmer temperatures increase (and cooler temperatures decrease) both accumulation and the amount of melt. In both cases the effect diminishes the melt-temperature relationship. Extracting the effect of changing accumulation rates from our record would therefore serve to strengthen rather than diminish its climatic signal. Furthermore, a study of ice cores and ice flow models¹⁷ has indicated that changes of the accumulation rate in Greenland and on the Devon and Agassiz Ice Caps have been generally <10% since at least 6 kyr BP (before present). These changes are small compared to the change in P from 5% at present to 100% in the early Holocene.

Because the ice cap is only 127 m thick at the drill sites and the accumulation rate is only 9.8 cm of ice per year, the time resolution at the beginning of our record is very coarse (Fig. 1). We can therefore draw only very broad conclusions about the climatic change. There are, however, four significant features in the record: (1) very high melt from 9.5 to 8.5 kyr; (2) a trend of decreasing melt from 8.5 to ~2.5 kyr; (3) a period of very low melt from 2.5 to 0.1 kyr; (4) the past 100 summers have included the warmest ones for over 1,000 years. As the present warm period follows the coldest period in the entire Holocene (Fig. 1a) it seems unusually warm if taken out of its 10,000-year context.

During the high-melt-period section from 9.5 to 8.5 kyr, melt layers often form 100% of the annual layer, that is, the entire annual layer melted. Part of the melt may have refrozen *in situ*, but part must have formed run-off, so that there was probably a negative balance at the drill site in some years. In this case melt would have been more than 100%. The complete melting and refreezing of the accumulation at these sites is approximately equivalent to the formation of a 10-cm ice layer in the snow-pack each year. This is 8.5 cm thicker than the annual ice layer forming over the past 50 years. At this latitude, an increase of melting of this amount would occur 200 m below the drill-site elevation¹⁸. Using a lapse rate of 1 °C per 100 m this translates into a summer cooling between 8.5 kyr BP and today of 2.0 °C.

There are, however, two local indirect climatic components in the melt-layer record. First, lowering of the surface during periods of negative balance produces a positive (warming) feedback. Second, isostatic recovery of the bed under the ice during the Holocene, which is of the order of 100 m in this area^{19,20},

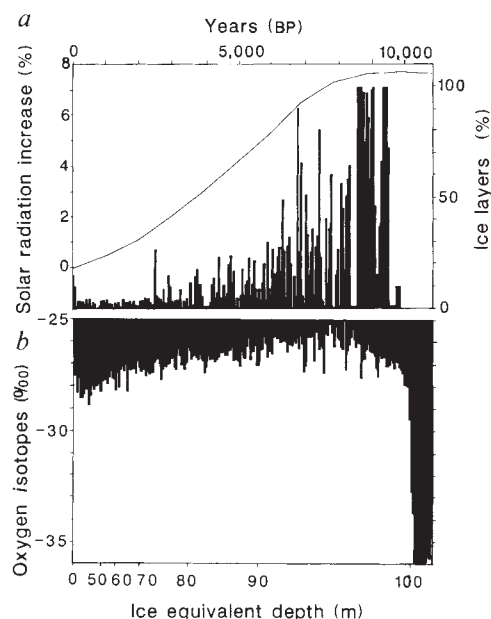


FIG. 1 a, Melt layers expressed as percentages of 50-year sections of the A-84 ice core. The upper curve represents increased solar radiation in June, July and August for the Northern Hemisphere due to changing orbital parameters (obliquity of the Earth's axis and precession of the equinoxes) relative to the present values²⁹. The timescale is based on: (1) measured thicknesses of sets of annual layers at various depths⁹; (2) comparison of major volcanic (acid) layers in the cores with those in the absolutely dated Dye-3 core in Greenland^{30,31}; (3) assigning a date of 10.75 kyr to the abrupt change in the stable-isotope ratios at the beginning of the Holocene period³². Note that the abrupt change of melt at the beginning of the record does not coincide with the rapid change in $\delta^{18}\text{O}$ values (lower profile). This may be due to melt features being missed in this part of a single core where fracturing, possibly due to a change of ice properties across the Holocene/Pleistocene transition, obscured the stratigraphy but did not affect the isotope analyses. b, Stable-isotope ratios ($\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{standard}}] - 1$, where the standard is SMOW, standard mean ocean water) for 50-yr sections of the ice core.

introduces a negative (cooling) feedback. The part played by each effect is difficult to assess. Based on the evidence from the late Wisconsin ice 12.5 m above the bed, we suspect that the thinning effect was small. This ice would have been very much thicker 9 kyr ago, as it has thinned since then because of ice flow away from the site. But the total ice thickness is only 127 m and this does not allow for very much thinning. On the other hand, 100 m of isostatic recovery over the past 8,000 to 9,000 years could contribute ~40% of the cooling shown over the same period in Fig. 1a.

Figure 1b shows, however, that there is a similar Holocene cooling trend from 9.5 kyr to the present in the stable-isotope record, which is an annual record rather than a seasonal one. The relationship between the stable-isotope data and temperature has already been shown to be independent of changing elevation in our area of study¹⁸. Assuming a relationship of 0.62‰ per 1 °C (ref. 21), we calculate a Holocene cooling from 9.5 kyr to the present of ~2.5 °C (Fig. 1b). This is similar to the summer value of 2.0 °C. The similarity between these two values indicates the summer melt-layer distribution is largely a direct climatic record.

There is a close relationship between overall glacier mass balance and melting in the firn zone of an ice cap as represented by any ice-core melt record³. Canadian high-Arctic ice caps and glaciers have slightly negative balances today²². So, if we assume that the present level of melt at the drill site (Fig. 1a) is representative of a slightly negative balance for the ice cap in general, the melt record indicates that the ice cap suffered a highly negative balance in the early- and mid-Holocene but a persistently positive balance from ~3 kyr to 0.1 kyr BP.

The case for glacier retreat (negative balance) in the high Arctic in the early Holocene is supported by some of the glacial geology records⁶⁻⁸, which indicate general glacier retreat beginning at, or before, 9–10 kyr BP. There is also evidence from nearby Greenland²³, Northern Ellesmere²⁴ and Axel Heiberg²⁵ for glacier growth during the second half of the Holocene. Our melt record is, however, in serious disagreement with the very detailed work of England²⁶ done in the area immediately to the north of Agassiz Ice Cap, which indicates slow glacier retreat beginning 8 kyr ago with rapid retreat beginning only after 6.2 kyr BP.

The increase of summer radiation owing to the change in the orbital parameters for the Northern Hemisphere is also shown in Fig. 1a; there seems to be a good correlation between the two. Although radiation does play an important part in the high-Arctic summer energy balance^{27,28}, we find no evidence for melt layers in our cores representing the late Wisconsin, when summer solar radiation values in the high northern latitudes were equally high. Thus the relationship between ice melt and the orbital parameters cannot be as direct as the record of the past 10,000 years indicates. □

19. England, J. *Quat. Res.* **9**, 185–202 (1976).
20. Blake, W. *Radiocarbon Dates* Vols. 26, 34 (Geol. Surv. Can. Ottawa, 1987).
21. Dansgaard, W., Johnsen, S. J., Clausen, H. B. & Gundestrup, N. *Meddelelser om Grønland* Bd 197, Nr. 2, 1–53 (Reitzels, Copenhagen, 1973).
22. Fisher, D. A. & Koerner, R. M. in *Climatic Change in Canada* Vol. 3 (ed. Harington, C. R.) 50–69 (Natn. Mus. Nat. Sci., Ottawa, 1983).
23. Kelly, M. *Geol. Surv. Greenland Rep.* No. 96, (1980).
24. Hattersley-Smith, G. in *Climatic Changes in Arctic Areas During the Last Ten Thousand Years* (eds Vasari, Y., Hyvärinen, H. & Hicks, S.) 137–148 (Oulu University, Finland, 1972).
25. Muller, F. in *Proc. Symp. Arctic Heat Budget Atmospheric Circulation* (ed. Fletcher, J. O.) 136–156 (Rand Corp., Santa Monica, 1966).
26. England, J. *Can. J. Earth Sci.* **20**, 895–917 (1983).
27. Havens, J. *Meteorology, Res. Rep.* No. 2 (McGill University, Montreal 1964).
28. Paterson, W. S. B. *The Physics of Glaciers* (Pergamon, Oxford, 1981).
29. Kutzbach, J. E. in *The Geology of North America* Vol. K-3 (eds Ruddiman, W. F. & Wright, H. E. Jr) 425–446 (Geol. Soc. Am., Boulder, 1987).
30. Hammer, C. U. *et al. J. Glaciol.* **20**, 3–26 (1978).
31. Hammer, C. U. *Jökull* **34**, 51–65 (1984).
32. Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284–291 (1986).

ACKNOWLEDGEMENTS. We thank N. Reeh for drawing our attention to the effect of changing elevation on our records. We also thank the Geological Survey of Canada for logistic support and the Department of Glaciology, University of Copenhagen, for the oxygen isotope analyses.

Detection of a volcanic fracture opening in Japan using Global Positioning System measurements

Seichi Shimada*, Yukio Fujinawa, Shoji Sekiguchi, Shiro Ohmi, Takao Eguchi & Yoshimitsu Okada

National Research Center for Disaster Prevention, Tennodai 3-1, Tsukuba, Ibaraki 305, Japan

THE Global Positioning System (GPS) is a powerful tool for detecting crustal deformation¹; for example, the technique has been used to detect plate boundary movements in California². Here we report observations of crustal deformation associated with seismic swarm and sea-floor volcanic activity off the east coast of the Izu Peninsula, Japan, in July 1989, using a fixed-point GPS network. These measurements have enabled us to capture some rarely observed features of seismic and volcanic activity. We have, for the first time, used GPS fixed-point measurements to follow the evolution with time of the crustal movements; such measurements provide a continuous uninterrupted record of deformation. Our observations were supported by independent data from other sources, thus providing further evidence for the utility of the GPS.

Since 1978, seismic swarms have occurred frequently in an area of radius ≤ 20 km off the east coast of the Izu Peninsula³. The maximum uplift, ~30 cm, was measured in an area (~40 × 20 km) to the west of the swarm region between 1978 and 1988 (ref. 4). The largest earthquake in this area (Richter scale magnitude = 6.7) occurred in 1980 and the most recent swarm before 1989 occurred in July and August 1988 (the magnitude of its largest event was 5.2 (ref. 5)).

Around 20 May 1989, a small-scale seismic swarm occurred for several days in this area. After one month's quiescence, an extensive swarm began again with peak activity between 4 and 10 July, during which (9 July) the largest (magnitude 5.5) earthquake occurred (strike-slip mechanism with slight reverse-slip component and maximum compressional axis of NNW–SSE). Volcanic tremors then occurred on 11 and 12 July, followed by a submarine volcanic eruption on 13 July (ref. 6), later named the Teishi volcano. Figure 1 shows the swarm epicentral region and the volcano.

Since March 1988, we have used a fixed-point GPS network (Fig. 1) to track GPS satellites daily and observe crustal movements in central Japan for the purpose of earthquake-prediction research⁷. The network overlaps the monitoring networks of

Received 4 August; accepted 28 December 1989.

1. Ahlmann, H. W. *Glaciological Research on the North Atlantic Coasts*, R. geogr. Soc. Res. Ser. No. 1 (London, 1948).
2. Hattersley-Smith, G. *Geogr. Annl.* **45**, 139–141 (1963).
3. Koerner, R. M. *Science* **196**, 15–18 (1977).
4. Herron, M. M., Herron, S. L. & Langway, C. C. *Nature* **293**, 389–391 (1981).
5. Hübner, W. D. & Langway, C. C. in *Polar Oceans* (ed. Dunbar, M. J.) 589–602, (Arctic Inst. N. Am. Montreal, 1977).
6. Bednarski, J. *Can. J. Earth Sci.* **23**, 1343–1355 (1986).
7. Lemmen, D. S. *Can. J. Earth Sci.* (in the press).
8. Blake, W. S. Jr *Geogr. Annl.* **57**, 1–71 (1975).
9. Fisher, D. A. *et al. Nature* **301**, 205–209 (1983).
10. Koerner, R. M., Fisher, D. A. & Paterson, W. S. B. *Can. J. Earth Sci.* **24**, 296–301 (1987).
11. Fisher, D. A. & Koerner, R. M. *J. Glaciol.* **32**, 501–510 (1986).
12. Koerner, R. M. *Science* **244**, 964–968 (1989).
13. Robin, G. de Q. & Johnsen, S. J. in *The Climatic Record in Polar Ice Sheets* (ed. Robin, G. de Q.) 47–52 (Cambridge University Press, 1983).
14. Raisbeck, G. M. *et al. Nature* **326**, 273–277 (1987).
15. Herron, M. M. & Langway, C. C. Jr in *Greenland Ice Core: Geophysics, Geochemistry, and the Environment* (eds Langway, C. C., Oeschger, H. & Dansgaard, W.) 77–84 (Am. Geophys. Un., Washington, DC, 1985).
16. Clausen, H. B. *et al. Ann. Glaciol.* **10**, 10–15 (1988).
17. Paterson, W. S. B. & Waddington, E. D. *Rev. Geophys. Space Phys.* **22**, 123–130 (1984).
18. Koerner, R. M. *J. Glaciol.* **22**, 25–41 (1979).

* Present address: Institute of Geophysics and Planetary Physics, A-025, Scripps Institution of Oceanography, La Jolla, California 92093-0225, USA.

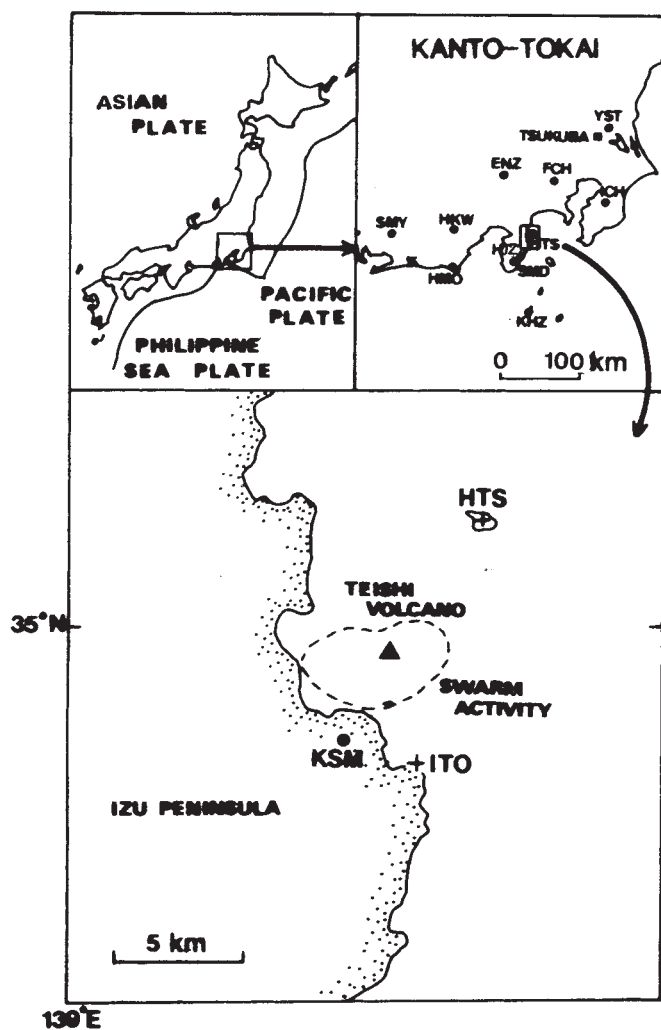


FIG. 1 Top, locations of Kanto-Tokai fixed-point GPS sites (closed circle) and volumetric strain meter at the HIZ site (triangle). Bottom, locations of fixed-point GPS sites, HTS and ITO (+), daily EDM site at KSM (●), the seismic-swarm epicentral region in July 1989 (dashed line), and the Teishi sea-floor volcano which erupted on 13 July 1989 (▲).

micro-earthquake and crustal tilt in the same region⁸. A Mini-Mac 2816AT GPS receiver was installed at each of the ten stations including the Hatsushima (HTS) site, which is located to the north-east of the swarm region. The Mini-Mac 2816AT is a dual-frequency (L1 at 1575.42 MHz and L2 at 1227.60 MHz) carrier-beat phase receiver and is able to track satellites every day automatically. We also installed a GPS receiver of the same type at the Ito (ITO) site in the upheaval area in January 1989 (Fig. 1). Unfortunately there are obstacles towards the west and south at the ITO site, so that we could not determine the position of this site until a new satellite became available in mid-May.

Here we discuss the GPS phase data from the HTS and ITO sites obtained in the period between May and September 1989, which includes the period of swarm and eruption activities. We exclude the other eight sites because the deformation seems to have resulted in a relative displacement between these two sites only. The data are collected over a period of 4 h with 30-s sampling and four-satellite (satellite PRN No. 3, No. 11, No. 13 and No. 14) tracking. We used broadcast ephemerides to calculate satellite orbits. Figure 2 shows the preliminary results of time variations of the HTS-ITO baseline vector. Owing to severe ionospheric activity and unfavourable satellite conditions over the sites, the analysis of the data obtained on certain days is very difficult, and we have not yet completed the analysis of all the data. Furthermore, we could not obtain the baseline solution for the last four days of the peak swarm activity in July, because the broadcast ephemerides of the satellite PRN No. 14 were not available.

We found significant crustal deformation in the period of the peak swarm activity in July. The difference between the average values of thirteen data before the swarm and ten data after the swarm indicates that the ITO site moved 13.6 cm to the south, 6.2 cm to the west, and ~5 cm up relative to the HTS site, and the HTS-ITO baseline length extended 14.5 cm. The error bars of each single plot are ~3 mm for horizontal components and baseline length, and ~10 mm for the vertical component. The r.m.s. scatter of the thirteen data points before the swarm are 3 mm, 4 mm and 13 mm for the latitudinal, longitudinal and vertical components respectively, and 3 mm for the baseline length respectively. The larger scatters after the swarm originate from the severe ionospheric activity. In the series of swarm and volcanic activities during May and July 1989, obvious

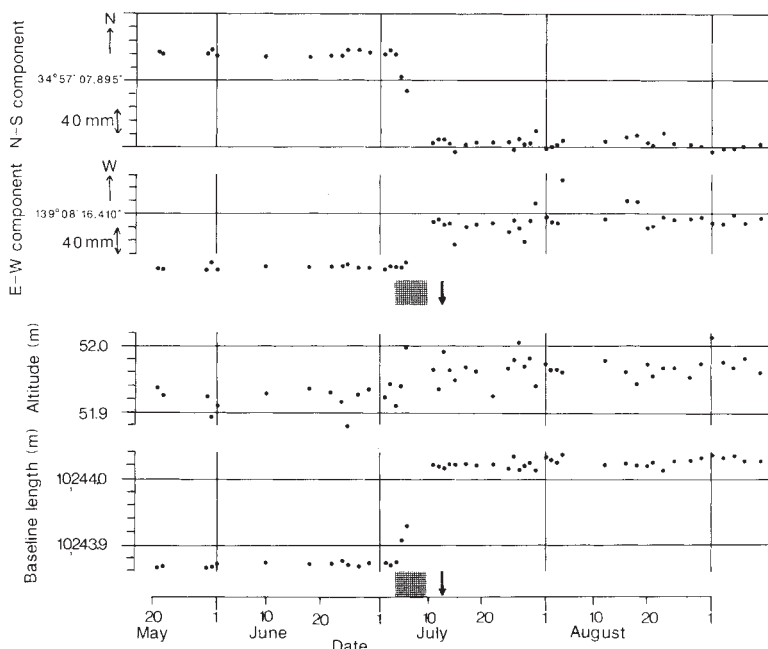


FIG. 2 Preliminary results of time variations of the position of the ITO site referenced to that of the HTS site. *a*, Latitudinal, *b*, longitudinal and *c*, vertical vector components of the ITO site, and *d*, the baseline length, are shown. The shaded region between 4 and 10 July indicates the peak activity of swarm, and the arrow denotes the sea-floor eruption on 13 July 1989.

crustal movements occurred only in the period of the peak swarm activity in July, and the movements ceased before the occurrence of the volcanic tremor and submarine eruption.

In the eastern part of Izu Peninsula, several continuous crustal-movement observations were also made, using, for example, a borehole tilt meter at the ITO site (National Research Center for Disaster Prevention), a borehole volumetric strain meter at the Higashi-Izu (HIZ) site ~30 km south-west of the swarm region (Japan Meteorological Agency), and daily EDM (electronic distance measuring) measurements between the HTS and Kusumi (KSM) sites ~3 km west of the ITO site (Earthquake Research Institute, University of Tokyo) (Fig. 1). The crustal movements observed by these instruments were reported at the Japanese Coordinating Committee for Earthquake Prediction on 7 August 1989. The tilt meter detected ESE-down crustal tilt of ~20 μ rad during the peak swarm, whereas the movements were <5 μ rad before and after the swarm. The strain meter at the HIZ site showed a crustal compressive strain of 10^{-6} and the EDM baseline length extended ~23 cm during the peak-activity period.

The crustal movements detected by the GPS measurements indicate a fracture opening around the swarm region. The swarm activity is interpreted as the fracture of crust caused by the opening accompanied with the upward movements of magma. From the epicentral distribution of the swarm, the fracture can be assumed to extend in a E-W to NW-SE direction and to

include Teishi volcano⁹. The direction is almost perpendicular to the predominant deformation of the GPS baseline, indicating a tensile movement. The crustal-tilt change at the ITO site, the crustal compression at the HIZ site and the EDM baseline extension are also consistent with the tensile model. The observed time variations of the GPS baseline reveal that the opening was completed before the volcanic tremors and submarine eruption.

Because the volcanic tremor is thought to originate from magma movements, the fact that the volcanic tremors and the sea-floor volcanic eruption occurred after the tensile deformation indicates that the magma movements that caused these volcanic phenomena were confined to a very limited surface layer. These results demonstrate that the GPS fixed-point observation can reveal the time sequence of the crustal movements associated with seismic swarm and volcanic activity. $\square$

Received 11 September; accepted 15 December 1989.

1. Prescott, W. H. *et al. Science* **244**, 1337–1339 (1989).
2. Davis, J. L. *et al. J. geophys. Res.* **94**, 13635–13650 (1989).
3. Ishida, M. *Proc. Earthquake Prediction Res. Symp.* 51–60 (1987) (in Japanese).
4. Geographical Survey Institute Rep. Coordinating Com. Earthquake Prediction **41**, 283–302 (1989) (in Japanese).
5. Japan Meteorological Agency, Rep. Coordinating Com. Earthquake Prediction **41**, 227–242 (1989) (in Japanese).
6. Swinbanks, D. *Nature* **340**, 175 (1989).
7. Shimada, S. *et al. J. geod. Soc. Japan* **35**, 85–95 (1989).
8. Okada, Y. *Nature* **312**, 500–501 (1984).
9. Okada, Y. & Yamamoto, E. *Prog. Abstr. Seismol. Soc. Japan* **2**, 118 (1989) (in Japanese).

Possible electrokinetic origin of large magnetic variations at La Fournaise volcano

J. Zlotnicki & J. L. Le Mouél.

Institut de Physique du Globe, Observatoires Volcanologiques et Laboratoire de Géomagnétisme, 4 place Jussieu, 75252 Paris Cedex 05, France

ABNORMAL time variations of the geomagnetic field related to volcanic activity have been reported for several decades^{1–10}. Two main candidates have been considered for the source mechanism of these volcanomagnetic variations: piezomagnetism and electrokinetic effects. We considered both of these mechanisms in our attempts to interpret the most spectacular volcanomagnetic variations observed on Piton de la Fournaise volcano (Réunion Island)¹¹, but now new analysis of the latest data, which covers a time span of more than three years, leads us to think that water circulation is the main agent of the strong magnetic signals observed on the volcano. In addition to the signals that we correlated with the volcano activity^{11,12}, longer-period components also exist, as shown here, with amplitudes and characteristics that have not been reported before. We propose that these magnetic variations are generated by electrokinetic effects, in which case the water circulation system into this massif would be a sensor for detecting stress changes linked to the volcano activity.

Up to seven stations P_i have been installed since 1985 on the volcano (Fig. 1). The magnetic field intensity $B(P_i, t_k)$ is measured with a sampling interval $t_{k+1} - t_k$ of one minute. To eliminate the main part of the time-varying transient field—which is due to ionospheric and magnetospheric currents—we calculate the differences $\Delta B(P_i, P_j, t_k) = B(P_i, t_k) - B(P_j, t_k)$ using the CSR station (Fig. 1) as the reference station P_1 . Here we will consider only daily mean values.

Figure 2a,b shows some of the differences $\Delta B(P_i, P_j, t)$ (called hereafter ΔB_{ij}) in the years 1986–1988. A clear strong annual wave appears on the curves ΔB_{31} , ΔB_{51} , ΔB_{61} , with a peak-to-peak amplitude reaching ~20 nT for ΔB_{31} (the difference (PMC–

CSR)). This annual wave is also observed on ΔB_{21} with a smaller amplitude. From several data recorded at P_7 (which is 15 km from the volcano, so we can assume that there is no abnormal effect), it can be shown that the annual wave on ΔB_{71} is small (<2 nT). The annual wave on ΔB_{41} (P_4 is close to the unwelded sommital calderas) could reach 10 nT peak to peak, but is difficult to estimate because of strong variations linked with eruptions and a long-term drift. A roughly linear trend also appears on the ΔB_{51} curve with a decrease of the order of 30 nT in 3 years.

There are several possible causes of those variations that we can discount. First, they are not due to drifts in the magnetometers (the quartz frequency is regularly checked) or to magnetic variations arising from sensor displacement (a tilt variation of 100 μ rad would generate an anomaly of less than a few nT at the sensor level (3 m high) in a local gradient of 100 nT m⁻¹)^{11,13–15}. Second, the annual variation of the temperature at ground level is 6 °C, which could induce, through the magnetization of the uppermost layers of the volcano (a few metres), an anomalous annual magnetic field. We can rule out this possibility using order-of-magnitude calculations (the temperature coefficients of the magnetizations are $<5 \times 10^{-4}$ °C⁻¹ around 20 °C) (J. P. Pozzi, personal communication); this also applies to variations arising from the cooling of the volcano by seasonal rainfall. Third, for the long-term drift, magnetization could be acquired as the volcano cools, but this is not in agreement with the observed short-time variations^{11,12}. Fourth, piezomagnetic effects are too weak to explain large annual variations and a long-term drift of ~30 nT. Fifth, we can envision magnetic induction effects¹⁶. Absolute field measurements on the array and on the PDB amagnetic pillar (Fig. 1) have shown that the anomalous field generated by the magnetized rocks is everywhere <3,000 nT and the mean normal field is 38,000 nT. Then if $|\Delta b|$ is the amplitude of the external ionospheric annual variation, the amplitude of the magnetic induction effect is at most $2|\Delta b|$ (3,000/38,000). But $|\Delta b|$ is only of the order of 10 nT at the Hermanus and Antananarivo observatories^{17,18} and at P_1 station (on the intensity; Fig. 3). The same reasoning shows that the observed long-term drifts cannot be due to a magnetic induction effect related to the secular variation of the core-generated field (a few tens of nT per yr in Réunion). Sixth, there

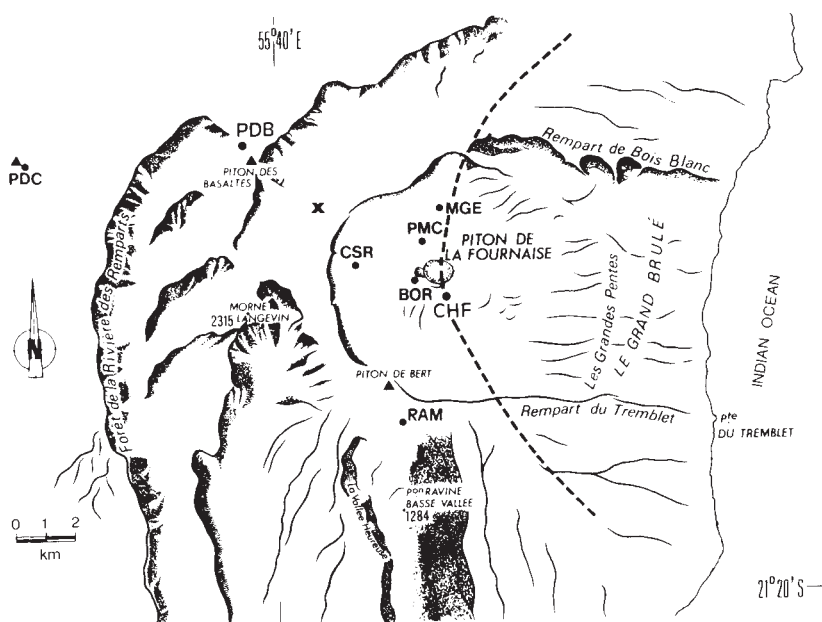


FIG. 1 Schematic map of Piton de la Fournaise volcano. ●, Magnetic stations; the numbering of the stations in the main text are 1 (CSR), 2 (RAM), 3 (PMC), 4 (BOR), 5 (MGE), 6 (CHF), 7 (PDC); PDB is an amagnetic pillar; ×, the climatological station; ----, the main fracture zone.

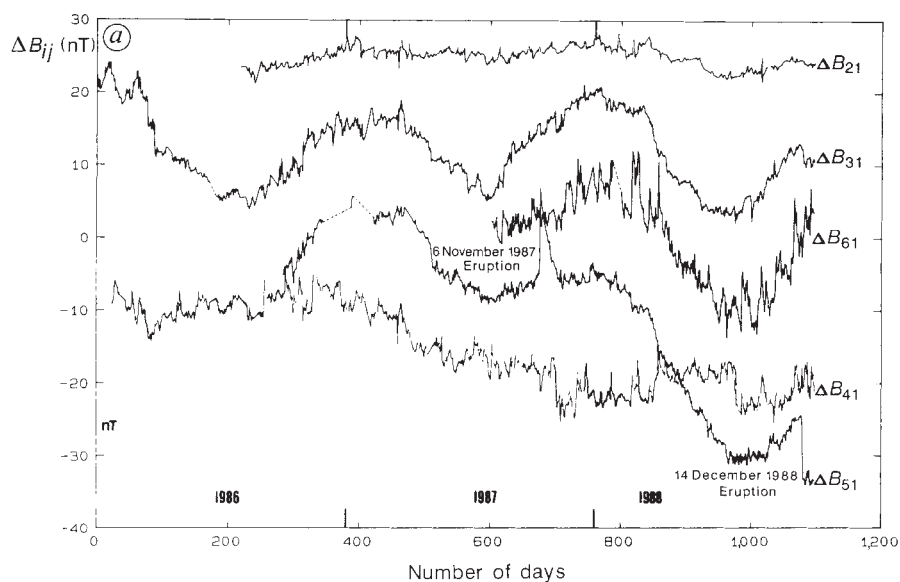
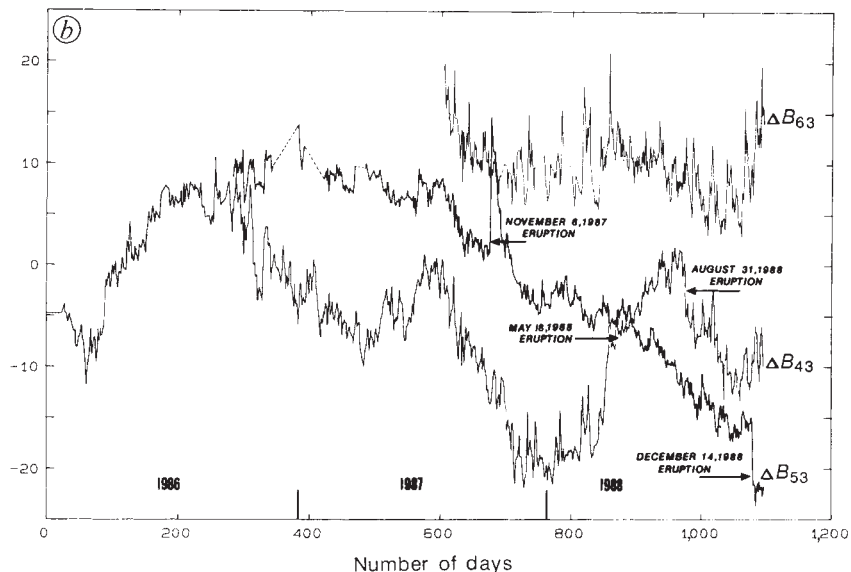


FIG. 2 a, Daily means of ΔB_{21} . b, Daily means of some other ΔB_{ij} .



could be an abnormal magnetic field due to electric currents induced through conductivity heterogeneities inside the volcano by the transient external field. A magnetic field of this kind does exist for shorter periods: $\Delta B(P_i, P_j, t_k)$ differences present a daily wave with an amplitude that can reach 30% of the amplitude of the 40-nT daily solar variation¹¹. But the ratio of this abnormal secondary field to the inducing normal field varies grossly as $1/T$, where T is the period of the considered variation¹⁹, and the amplitude of the annual variation is not larger than 10 nT, as mentioned above.

We therefore think that we are observing an electrokinetic magnetic anomaly^{20,21} (fluid motion through a porous medium induces an electric current that is the result of the charge separation between the fixed ions in the matrix and the mobile ions (of opposite sign) in the liquid along the pore walls). It rains heavily on the water-saturated, fractured, unwelded massif (up to 6 m yr^{-1}) and there is an annual variation in the rainfall (Fig. 3). Furthermore, the most intense annual signals are observed at P_3 (PMC), P_5 (MGE) and P_6 (CHF) whereas this signal is weak at P_1 (CSR), P_2 (RAM), PDB pillar, and P_7 (PDC). The first three stations are located close to the well defined fracture zone (sometimes called the rift zone), a faulted vertical structure that cuts across the whole volcano²² (Fig. 1), and which is likely to play an important part in the water circulation through the massif.

Although little is known about the hydrology of the volcano²³ and the values of the parameters relevant to the electrokinetic theory, we can try to calculate whether this mechanism can give magnetic signals of the right order of magnitude. We suppose that the magnetic signal is due to an electric current circulating through the fractures and we approximate this fracture zone by a horizontal prism, infinitely long, with a rectangular cross-section 1 km wide and 1 km deep²² (these parameters are not critical). We assume that the pore pressure gradient in this porous structure, ∇P , is uniform and parallel to the trend of the structure. An electric-current density J is generated in the porous medium inside the fractures zone²¹:

$$J = -L_{12} \nabla P = (\epsilon \zeta / \eta) (\sigma_r / \sigma_f) \nabla P \quad (1)$$

Here ϵ is the dielectric constant of the fluid, ζ the zeta potential, η the viscosity of the fluid, and σ_r and σ_f are respectively the rock and fluid conductivity. $C = -L_{12} / \sigma_r = -(\epsilon \zeta / \eta \sigma_f)$ is called the cross-coupling or streaming coefficient and is expressed in V Pa^{-1} (refs 20, 21, 24, 25 and B. Nourbehecht and T. Madden, unpublished results). The different parameters entering equation (1) are poorly known. We use the following values, which were used in the cited references for similar situations: $\epsilon = 80 / (36\pi \times 10^9) \text{ F m}^{-2}$, $\eta = 10^{-4} \text{ N s m}^{-2}$, $\sigma_r / \sigma_f = 10^{-2}$, $\zeta = 10^{-2} \text{ V}$ (that is $C = 10^{-6} \text{ V Pa}^{-1}$ for $\sigma_f = 10^{-1} \text{ S m}^{-1}$). Substituting these in equation (1) we obtain $J = 10^{-9} \nabla P$. J is in A m^{-2} if ∇P is in Pa m^{-1} . Using the hydrostatic pressure gradient at depth due to the volcano topography for ∇P ($|\nabla P| \approx 1,000 \text{ bars} / 5 \text{ km} \approx 2 \times 10^4 \text{ Pa m}^{-1}$), then $J \approx 10^{-5} \text{ A m}^{-2}$, which gives a magnetic field of the order of 10 nT at the ground surface.

There seems to be no contradiction between the proposed mechanism and our observations, but the simplicity of our model and the rather arbitrary choice of the values of the physical parameters prevent any strong statement. We have considered the pressure gradient as being due to the volcano topography; one could also postulate a forcing mechanism due to the ocean.

We have previously¹¹ cited many examples of magnetic variations related to the activity of this volcano; these signals generally preceded, by a few days or a few weeks, a seismic crisis or a volcanic eruption. In ref. 11 we attributed them either to piezomagnetism or to electrokinetic effects. Now that we have observed the annual variation, we conclude that the second mechanism is the important one. In the ΔB_{ij} data (Fig. 2a,b) the volcanomagnetic signals, with amplitudes up to 15 nT and periods up to a few weeks, appear as a perturbation of the smooth annual or long-term drift signals (for example, we

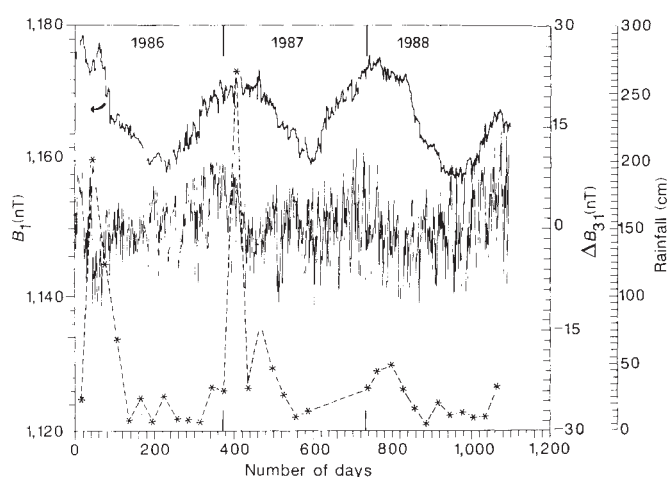


FIG. 3 Monthly means of the rainfall at the climatological station (dashed line). The upper curve is ΔB_{31} and the middle curve shows the daily values of B_1 after a parabolic trend has been removed. The amplitude of the annual ionospheric wave is $< 10 \text{ nT}$.

observed a sharp peak at the beginning of the 6 November 1987 eruption on the ΔB_{51} difference which has a life-time of a few tens of days¹². It is tempting to think that the mechanism producing all of the components of the observed signals should be the same: the electric-current density J circulating preferentially in the fracture zone, the intensity and space distribution which show an annual variation and a long-term drift, is also sporadically modified by the volcano activity (changes to either ΔP or porosity). The latter variations are probably generated by stress variations in the massif. The stress variations are in turn related directly to the deeper phenomena governing the volcano activity, such as changes in the overpressure in a magmatic reservoir. The stress variations in the massif are not large, probably smaller than a few tens of bars. Taking $L = 500 \text{ m}$ as a characteristic length scale of the stress variations distribution in this faulted and heterogeneous structure^{26,27}, the stress gradient due to the volcano activity may indeed be of the same order of magnitude as the pressure gradient responsible for the regular flow as estimated above.

Although our data are consistent with the electrokinetic mechanism, particularly when we consider the geometry of the anomalous field, more data are needed—especially electrical measurements. These are, however, difficult to obtain. In any case, these data should raise a renewed interest in magnetic and electric measurements in different contexts of tectonic activity. □

Received 14 July; accepted 27 December 1989.

- Rikitake, T. *Bull. Earthquake Res. Inst. Univ. Tokyo* **29**, 499–502 (1951).
- Yokoyama, I. *Bull. Earthquake Res. Inst. Univ. Tokyo* **35**, 567–572 (1957).
- Johnston, M. J. S. & Stacey, F. D. *J. geophys. Res.* **74**, 6541–6544 (1969).
- Davis, P. M., Jackson, D. B., Field, J. & Stacey, F. D. *Science* **180**, 73–74 (1973).
- Davis, P. M., Pierce, D. R., McPherson, R. L. & Dzursin, D. *Geophys. Res. Lett.* **11**, 233–236 (1984).
- Pozzi, J. P., Le Mouél, J. L., Rossignol, J. C. & Zlotnicki, J. *J. Volcan. geotherm. Res.* **5**, 217–237 (1979).
- Johnston, M. J. S., Mueller, R. J. & Dvorak, J. *Geol. Surv. Prof. Pap. U.S.* **1250**, 183–189 (1981).
- Zlotnicki, J. *J. Volcan. geotherm. Res.* **30**, 83–116 (1986).
- Rikitake, T. *J. Geomagn. Geoelectr., Kyoto* **39**, 47–61 (1987).
- Sasai, Y. et al. *J. Geomagn. Geoelectr., Kyoto* (in the press).
- Zlotnicki, J. & Le Mouél, J. L. *J. geophys. Res.* **90**, 9157–9171 (1988).
- Zlotnicki, J., Le Mouél, J. L. & Moreau, M. G. *Proc. Kagoshima Int. Conf. Volcanoes* 373–376 (1988).
- Le Mouél, J. L., Menvielle, M., Pozzi, J. P., Simon, G. & Zlotnicki, J. *J. Geomagn. Geoelectr., Kyoto* **36**, 609–619 (1984).
- Zlotnicki, J. et al. *J. Geomagn. Geoelectr., Kyoto* **38**, 151–164 (1986).
- Zlotnicki, J., Moreau, M. G. & Viodé, J. P. *J. Geomagn. Geoelectr., Kyoto* **39**, 487–500 (1987).
- Davis, P. M., Stacey, F. D., Zablocki, C. J. & Olson, J. V. *Phys. Earth planet. Inter.* **19**, 331–336 (1979).
- Chapman, S. & Bartels, S. *Geomagnetism* (Clarendon, Oxford, 1940).
- Courtilot, V. & Le Mouél, J. L. *J. geophys. Res.* **81**, 2941–2950, (1976).
- Le Mouél, J. L. & Menvielle, M. *Geophys. J. R. astr. Soc.* **68**, 575–587 (1982).
- Mizutani, H., Ishido, T., Yokokura, T. & Ohnishi, S. *Geophys. Res. Lett.* **3**, 365–368 (1976).
- Fitterman, D. V. *J. geophys. Res.* **83**, 5923–5928 (1978).

22. Bachelery, P. thesis, Univ. Clermont-Ferrand (1981).
23. Gourgand, B. & Stieljes, L. *Hydrogéologie* **2**, 95–107 (1988).
24. Nourbehecht, B. thesis, Mass. Inst. Technol. (1963).
25. Scheidegger, A. E. *The Physics of Flow Through Porous Media* 3rd Edn, 126–129 (University of Toronto Press, 1974).
26. Zlotnicki, J., Ruegg, J. C., Blum, P. A. & Bachelery, P. *J. Volcan. geotherm. Res.* (in the press).
27. Delorme, H. *J. Volcan. geotherm. Res.* **36**, 199–208 (1989).

ACKNOWLEDGEMENTS. We thank the members of Piton de la Fournaise Volcanic Observatory, especially J. C. Delmond and P. Tauchi. We also thank J. P. Pozzi for discussions, J. P. Poirier for comments, D. Gilbert for help and Dr Fitterman for suggestions.

Variations in effective elastic thickness of the North American lithosphere

Timothy D. Bechtel*, Donald W. Forsyth*,
Virgil L. Sharpton† & Richard A. F. Grieve‡

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

† Lunar and Planetary Institute, Houston, Texas 77058, USA

‡ Department of Energy, Mines and Resources, Ottawa, Ontario, Canada K1A 0Y3

THE isostatic response of the lithosphere to loading can often be successfully represented by a model in which the weight of topographic features on the Earth's surface is supported in part by stresses within a flexed elastic plate overlying an inviscid asthenosphere, and in part by the buoyancy forces associated with the deflection of density interfaces accompanying flexure^{1,2}. Previous estimates^{3–10} of the apparent thickness of the elastic plate in North America, based on the shapes of sedimentary basins or the deformation accompanying slip on faults, range from a few to more than a hundred kilometres—or equivalently, flexural rigidities ranging from about 10^{21} to 10^{25} N m⁻¹. Here we use a technique for estimating flexural rigidity that is not limited to sedimentary basins to map variations in the effective elastic thickness of the North American lithosphere. The effective elastic thickness ranges from a minimum of ~4 km in the Basin and Range Province to >100 km in the Precambrian core of the continent, supporting the idea that flexural rigidity increases with time since the last thermal event¹¹.

The technique used involves estimating the coherence between Bouguer gravity anomalies and topography as a function of wavelength. Bouguer gravity anomalies represent the attraction of lateral variations in sub-surface density and thus yield information about the buoyant compensation for surface topographic features. The technique is based on the fact that the relationship between topography and Bouguer anomalies as a function of wavelength is very different for surface and sub-surface loads that are applied to an elastic plate. At very long wavelengths, a surface topographic load can easily deflect an elastic plate, producing a compensating root, with its corresponding Bouguer gravity-anomaly signature. In this case, complete local (Airy) compensation occurs. A long-wavelength sub-surface load will also deflect the plate, creating surface topography that locally compensates the sub-surface load. As both surface and sub-surface loads are compensated in similar fashion (identical if there is only one sub-surface density interface and loads are applied at that depth), there will be a very high correlation or coherence between topography and Bouguer gravity anomalies at very long wavelengths.

At short wavelengths, loads are supported largely by stresses within the plate and the amplitude of the deflection of the plate is very small. At these wavelengths, surface topographic loads produce no significant Bouguer anomaly, and sub-surface loads produce gravity anomalies but no significant topographic signature. Thus, if surface and sub-surface loading are statistically independent processes, there will be no coherent relationship between topography and Bouguer anomalies at short

wavelengths. The wavelength of the transition between a coherent and an incoherent relationship is an indicator of the characteristic flexural wavelength of the elastic plate, which can be translated into the flexural rigidity of the plate. The longer the wavelength of the transition, the thicker or stiffer is the plate. The elastic thickness that best predicts the observed coherence is described as an effective or apparent elastic thickness, because the actual thickness of the lithosphere giving rise to the apparent elastic response may differ if parts of the lithosphere are yielding inelastically. The method of predicting the coherence for a given plate thickness and inverting for rigidity has been described fully^{12,13} and applied to data from East Africa^{14,15}, Australia¹⁶ and Asia¹⁷.

For the coherence analysis, original topography and gravity data compiled by the Gravity Data Center of the Geophysics Branch of Energy, Mines and Resources, Canada were averaged and interpolated onto a Lambert conformal projection with square 12-km grid spacing. We then re-interpolated these grid data onto grids within smaller rectangular regions using a gaussian weighting function with fall-off distance equal to the grid spacing. Figure 1 shows the observed coherence between topography and Bouguer anomalies as a function of wavelength for four such rectangular regions. In each case, there is a clear transition between high coherence at long wavelengths and low coherence at short wavelengths. The transition wavelength varies dramatically depending on the tectonic provinces included within the region. For example, in Fig. 1, the wavelength varies from ~100 km for region NBAR that covers primarily the northern Basin and Range (province outlines shown in Fig. 2) to ~800 km for region ECNS that includes mostly the eastern Canadian shield.

We calculated the coherence for more than 50 overlapping regions—some regions were designed to isolate, as much as possible, individual tectonic provinces, and others were designed to provide uniform sampling with rectangles of identical size and shape. Coherences were translated into elastic thickness using a model that assumes uniform thickness within an individual region. The minimum thickness found for any single rectangle was 4.9 km and the maximum was 123 km (for a rectangle centred in the Canadian shield). Non-uniform plate thickness within a region tends to broaden the transition between high and low coherence in a predictable manner (Fig. 2). If the thickness is non-uniform, the analysis assuming uniformity yields a best estimate of the effective elastic thickness that is a weighted average of the provinces in the region. Using broadened transitions as an indication of non-uniformity in a region, we estimate that the effective elastic thickness varies from a minimum of ~4 km in the Basin and Range to ~128 km in the core of the continent (Fig. 3). It is difficult to isolate completely these extreme values in an individual rectangle because of the irregular shape of geological provinces and because large regions are needed to capture the transition wavelength. In Fig. 3 we present contours summarizing the variability at logarithmic intervals (because the resolution of thickness is logarithmic). The particular interval is chosen because we are completely confident that a factor of two is significant: a province with an effective elastic thickness of 16 km can be distinguished without doubt from one with thickness of 8 or 32 km.

The elastic thicknesses found in this study differ dramatically from those reported in previous studies that used the statistical relationships between gravity and topography in North America. Earlier studies^{18,19}, using the admittance or ratio between gravity anomaly and topography as a function of wavelength rather than the coherence, found effective thicknesses ranging from zero in the western Cordillera to ~20 km in the eastern part of the continent, with an average <10 km. We have shown¹³ that the admittance technique is systematically biased towards low estimates when it is assumed that topography results from surface loading alone, as was done in the earlier studies. We also demonstrated that the coherence technique used here is not

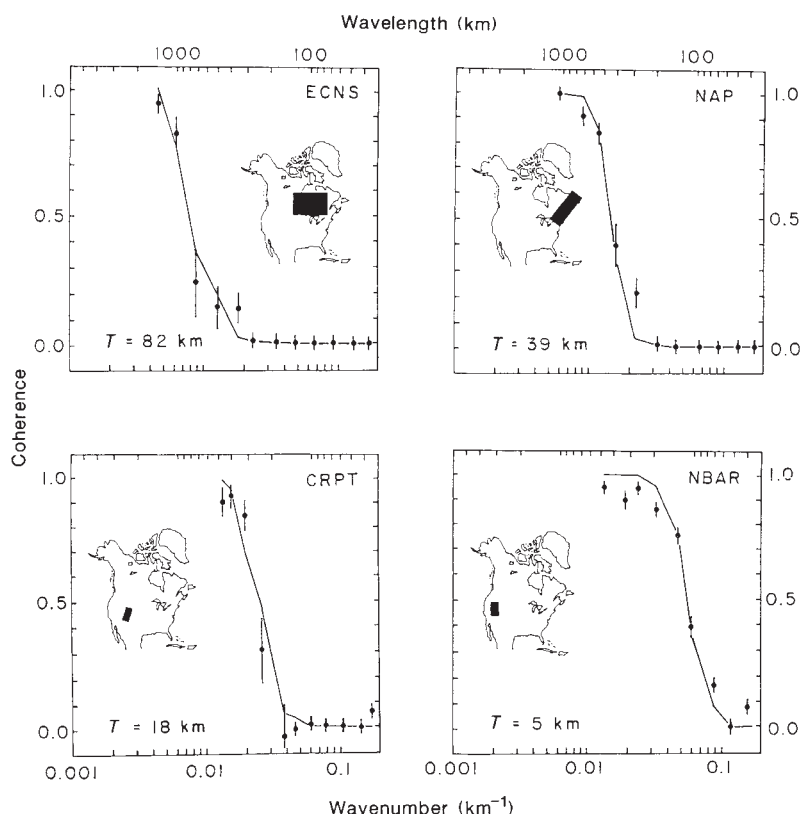


FIG. 1 Coherence as a function of wavelength for several example regions in North America. Shaded rectangles in index maps indicate areas covered: ECNS, eastern Canadian shield; NAP, northern Appalachians; CRPT, Colorado plateau; NBAR, northern Basin and Range. Vertical bars represent one standard error. Predicted coherence values are joined by solid lines. Coherence γ_0^2 defined as $\gamma_0^2 = C^2/E_0E_1$, with an unbiased estimate obtained from $\gamma^2 = (n\gamma_0^2 - 1)/(n - 1)$, where E_0 is the topography power spectrum, E_1 is the gravity power spectrum, C is the cross-spectrum, and all averages are over n independent Fourier coefficients in a waveband. Spectra calculated using two-dimensional, fast Fourier transforms (FFT) of regularly spaced points in rectangular map regions. Coherence was estimated for bands equally spaced in the wavenumber domain, although wavenumbers plotted are not equally spaced because the average wavenumber was weighted by amplitudes of individual Fourier components. Best-fitting elastic thickness T is indicated for each region. T is related to flexural rigidity D by $D = ET^3/12(1 - \nu^2)$, where E is Young's modulus and ν is Poisson's ratio, and to characteristic wavelength λ by $\lambda = 2\pi(D/\Delta\rho g)^{1/4}$, where g is the gravitational acceleration and $\Delta\rho$ is the density contrast at the interface that provides compensation for the applied load.

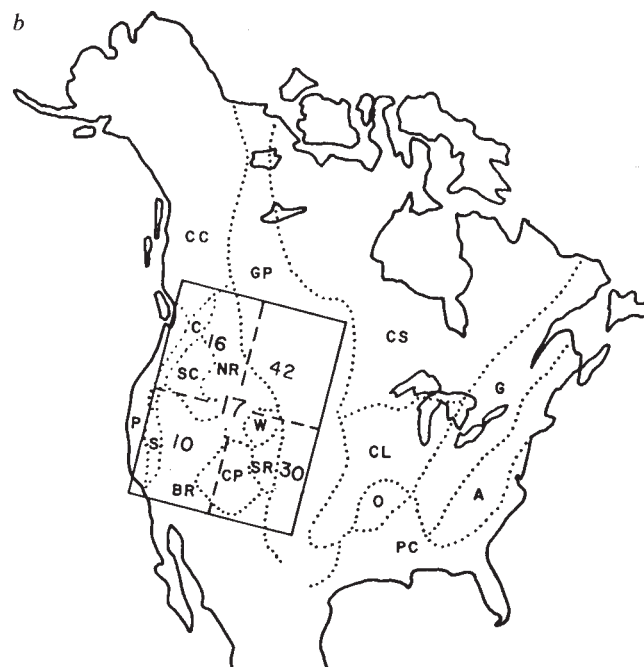
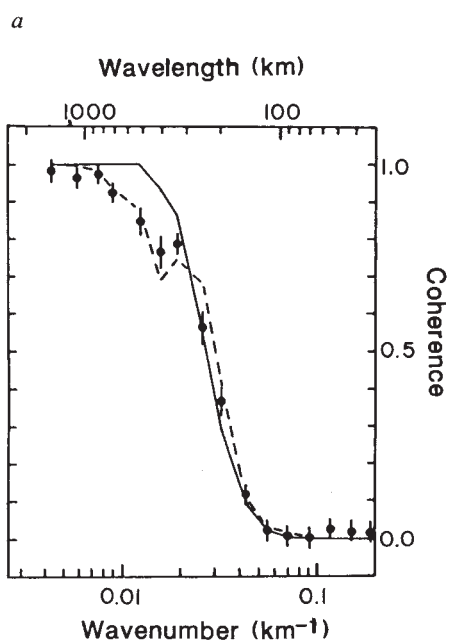


FIG. 2 Illustration of the effect of combining provinces with different effective elastic thicknesses into one region. *a*, shows the observed coherence for the large rectangular region in *b* that is outlined with a solid line and covers much of western North America. Assuming uniform plate thickness for the entire region, the best estimate of T is 17 km, but the predicted transition from high to low coherence is then narrower than observed (solid line). Dividing the large rectangle into four smaller regions yields estimates of T for the sub-regions ranging from 10 to 42 km. If the predicted coherences from each of the sub-regions are averaged, with weights for each sub-region proportional to the product of the power of the topography and gravity as

is done in weighting individual Fourier components in the usual coherence estimates, the predicted coherence (dashed line) successfully matches the broadened transition. Each of the four estimates of elastic thickness in the sub-regions are in themselves weighted averages, as the sub-regions cover more than one tectonic province. The extent of major tectonic provinces is indicated by dotted lines: A, Appalachian; BR, Basin and Range; C, Cascade; CC, Canadian Cordillera; CL, Central Lowlands; CP, Colorado Plateau; CS, Canadian Shield; G, Grenville; GP, Great Plains; NR, Northern Rockies; O, Ozark and Ouchita; P, Pacific Coast Ranges; PC, Coastal Plain; S, Sierra Nevada; SC, Columbia and Snake; SR, Southern Rockies; W, Wyoming Basin.

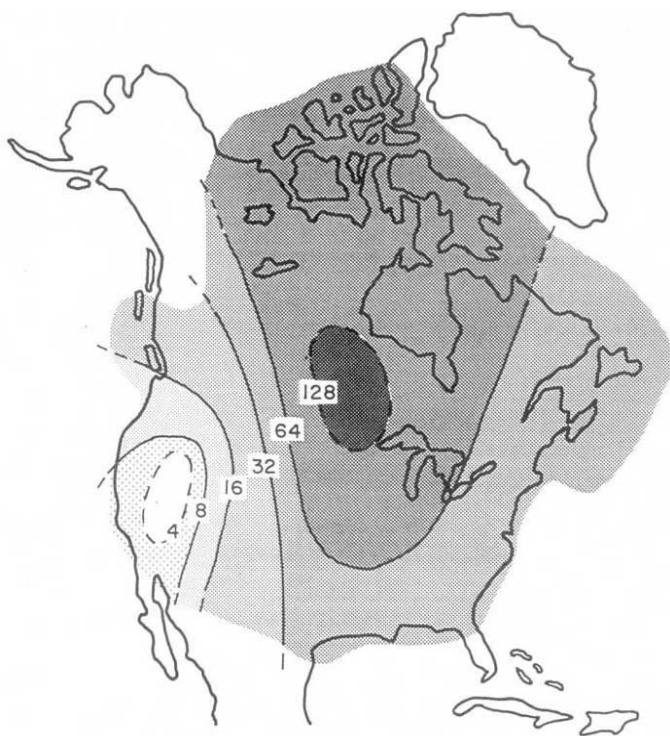


FIG. 3 Variations in effective elastic thickness of the North American lithosphere. Contour values labelled are in km, assuming Young's modulus $E = 1.0 \times 10^{11} \text{ N m}^{-2}$. Contours for 4 and 128 km are dashed to indicate that rigidity approaches or slightly exceeds these extreme values inside contours.

similarly biased. The average thickness that we find, several tens of kilometres, is consistent with that found in model studies of individual features³⁻¹⁰ and with an estimate based on the long-term erosional decay of topography²⁰.

The effective elastic thickness of the North American lithosphere (Fig. 3) is controlled largely by the thermal state of the lithosphere, with areas of reduced rigidity coinciding with high heat flow²¹⁻²³, high electrical conductivity of the mantle²⁴, low Pn (refracted P wave) velocities²⁵, and seismic travel-time delays^{26,27}. As pointed out by Karner *et al.*¹¹, the flexural rigidity increases with time since the last thermal event. The interior Precambrian shield regions have the highest rigidities, followed by the Palaeozoic Appalachian region in the east, and the lowest rigidities are found beneath the recently active Basin and Range province. There is also some indication that displacement on faults aids isostatic adjustment and reduces the apparent rigidity: experiments indicate that in western United States the apparent rigidity is azimuthally anisotropic with maximum rigidity parallel to the average orientation of faults in the region¹². Thus, the stiffness of individual, fault-bounded blocks such as the Sierra Nevada may be greater than the rigidity of the region as a whole^{19,28,29}. □

Received 2 October; accepted 20 December 1989.

- Barrell, J. J. *J. Geol.* **22**, 425-443 (1914).
- Gunn, R. J. *Franklin Inst.* **236**, 372-396 (1943).
- Cohen, T. J. & Meyer, R. P. *Geophys. Monogr.* **10**, 141-165 (1966).
- Walcott, R. I. *J. geophys. Res.* **75**, 3941-3954 (1970).
- Karner, G. D. & Watts, A. B. *J. geophys. Res.* **88**, 10449-10477 (1983).
- Stein, R. S. & Barrientos, S. E. *J. geophys. Res.* **90**, 11355-11366 (1985).
- Quinlan, G. M. & Beaumont, C. *Can. J. Earth Sci.* **21**, 973-996 (1984).
- Fulton, R. J. & Walcott, R. I. *Geol. Soc. Am. Mem.* **142**, 163-173 (1975).
- Sheffels, B. & McNutt, M. K. *J. geophys. Res.* **91**, 6419-6431 (1986).
- Bills, B. G. & May, G. M. *J. geophys. Res.* **92**, 11493-11508 (1987).
- Karner, G. D., Steckler, M. S. & Thorne, J. A. *Nature* **304**, 250-253 (1983).
- Bechtel, T. D. thesis, Brown Univ. (1989).
- Forsyth, D. W. *J. geophys. Res.* **90**, 12623-12632 (1985).
- Bechtel, T. D., Forsyth, D. W. & Swain, C. J. *Geophys. J.* **90**, 445-465 (1987).
- Ebinger, C. J., Bechtel, T. D., Forsyth, D. W. & Brown, C. O. *J. geophys. Res.* **94**, 2883-2901 (1989).
- Zuber, M. T., Bechtel, T. D. & Forsyth, D. W. *J. geophys. Res.* **94**, 9353-9367 (1989).
- McNutt, M. K., Diamant, M. & Kogan, M. G. *J. geophys. Res.* **93**, 8825-8838 (1988).

- Banks, R. J., Parker, R. L. & Huestis, S. P. *Geophys. J.* **51**, 431-452 (1977).
- McNutt, M. K. *J. geophys. Res.* **85**, 6377-6396 (1980).
- Stephenson, R. *Geophys. J.* **77**, 385-413 (1984).
- Jessop, A. M., Hobart, M. A. & Sclater, J. G. *Geotherm. Ser.* **5**, (1976).
- Pollack, H. N. & Chapman, D. S. *Tectonophysics* **38**, 279-296 (1977).
- Blackwell, D. D. *Geol. Soc. Am. Mem.* **152**, 175-208 (1978).
- Gough, D. I. *J. Geomag. Geoelect.* **26**, 105-123 (1974).
- Tucker, W., Herrin, E. & Friedman, H. W. *Bull. Seis. mol. Soc. Am.* **58**, 1243-1260 (1968).
- Dziewonski, A. M. & Anderson, D. L. *J. geophys. Res.* **88**, 3295-3314 (1983).
- Hales, A. L. & Doyle, H. A. *Geophys. J.* **13**, 403-415 (1967).
- Glazner, A. F. & Schubert, G. *J. geophys. Res.* **90**, 5405-5409 (1985).
- Chase, C. G. & Wallace, T. E. *J. geophys. Res.* **93**, 2795-2802 (1986).

ACKNOWLEDGEMENTS. This work was supported by the NSF.

Food thresholds and body size in cladocerans

Z. Maciej Gliwicz*

Max-Planck-Institut für Limnologie, Postfach 165, Plön, FRG

ACCORDING to the size-efficiency hypothesis¹, body size is one of the most important determinants of the relative abundance of planktonic animals in nature. First, large-bodied species are more vulnerable to predation by visually feeding planktivorous fishes. Second, they are superior competitors for resources, being able to grow and reproduce at lower food concentrations. But whereas the first assumption has been confirmed by many field and experimental data (for review see refs 2-4), the second suggestion lacks experimental evidence and has been frequently questioned⁵⁻¹². I now present experimental data from eight filter-feeding species (family *Daphnidae*) that strongly support the cornerstone assumption of the competitive aspect of the size-efficiency hypothesis. These data show that the threshold food concentration necessary to assure that assimilation equals respiration¹³, is lower for the large-bodied species than it is for the small-bodied species under steady-state and low-mortality conditions.

When two filter-feeding herbivore species compete for a single limiting food resource under steady-state conditions and in the absence of predation, the species with the lowest food threshold concentration should always win, regardless of its maximum intrinsic reproductive potential. In an environment with high predation and wide food-level fluctuations, a high intrinsic rate of population increase and the ability to withstand long periods of starvation may also be important. Under steady-state conditions and in the absence of predation, however, the species that can grow and reproduce at the lowest food concentration may keep the resource level below the threshold concentrations of other species for a time sufficient to cause their exclusion.

Assuming that the food threshold concentration is species-specific for a particular type of food in a common environment, I believe that the food threshold concentration is a good measure of competitive superiority for a given species relative to other filter-feeding species with the same feeding modes. The food threshold concentration is a reliable measure as long as mortality rates are low, and different species-specific reproduction rates have not become important in equalling mortality increases. In the case of high mortality, the threshold for a positive intrinsic rate of population increase would no longer be reflected by the threshold for individual growth (see ref. 14 for a review on resource competition in zooplankton).

To compare this value among cladoceran species of different body sizes, I selected eight *Daphnidae* species (Table 1) which do not seem to show selective sieving or the ability to discriminate between individual particles on a chemosensory basis^{15,16}. For each species, I determined body growth rates at different concentrations of a single food resource. I kept each concentra-

* Permanent address: Department of Hydrobiology, University of Warsaw, Nowy Swiat 67, 00-046 Warsaw, Poland.

TABLE 1 Body sizes and ages at maturation of eight daphnids

Daphnid	Mean body length* of:			Body carbon† (µg, mean ±1 s.d.) of:		Age (days) at which clutch of eggs is released into brood chamber ≥8
	neonate (mm) 1‡	2-day-old ≥60	6-day-old ≥80	2-day-old ≥6	6-day-old ≥8	
<i>D. magna</i>	0.87	1.83	2.38	11.4 ± 2.7	35.4 ± 9.3	6.1
<i>D. pulicaria</i>	0.72	1.17	1.80	4.4 ± 0.7	15.9 ± 1.7	6.2
<i>D. pulex</i>	0.69	1.14	1.47	4.3 ± 0.2	13.5 ± 2.7	5.9
<i>D. hyalina</i>	0.61	1.12	1.43	3.7 ± 0.2	9.0 ± 1.1	6.2
<i>D. galeata</i>	0.59	0.91	1.03	2.5 ± 0.5	7.0 ± 1.4	5.5
<i>D. ambigua</i>	0.42	0.60	0.91	0.9 ± 0.3	4.4 ± 1.0	5.5
<i>D. cucullata</i>	0.38	0.59	0.89	0.9 ± 0.1	3.1 ± 0.2	5.6
<i>C. reticulata</i>	0.32	0.51	0.59	1.1 ± 0.1	3.0 ± 0.1	5.0

Body sizes and ages at maturation of the eight daphnids species (seven *Daphnia* and one *Ceriodaphnia*) at high food (*S. acutus*) concentrations (0.20 to 0.25 mg C l⁻¹). At the lower food level (0.025 to 0.031 mg C l⁻¹) body length of 6-day-old individuals was only 20–30% reduced, whereas body carbon was 60–80% reduced. At the lowest level, no eggs were produced in any of the eight species.

* From the upper edge of the eye to the base of the tail spine.

† In each replicate, estimated in 5–25 animals dried to constant weight at 60 °C and combusted in a Pregl–Roth oven for organic carbon estimate on an UNDR infrared gas analyser²⁴.

‡ Minimum: the length of the smallest neonate found 2 days before starting experiments.

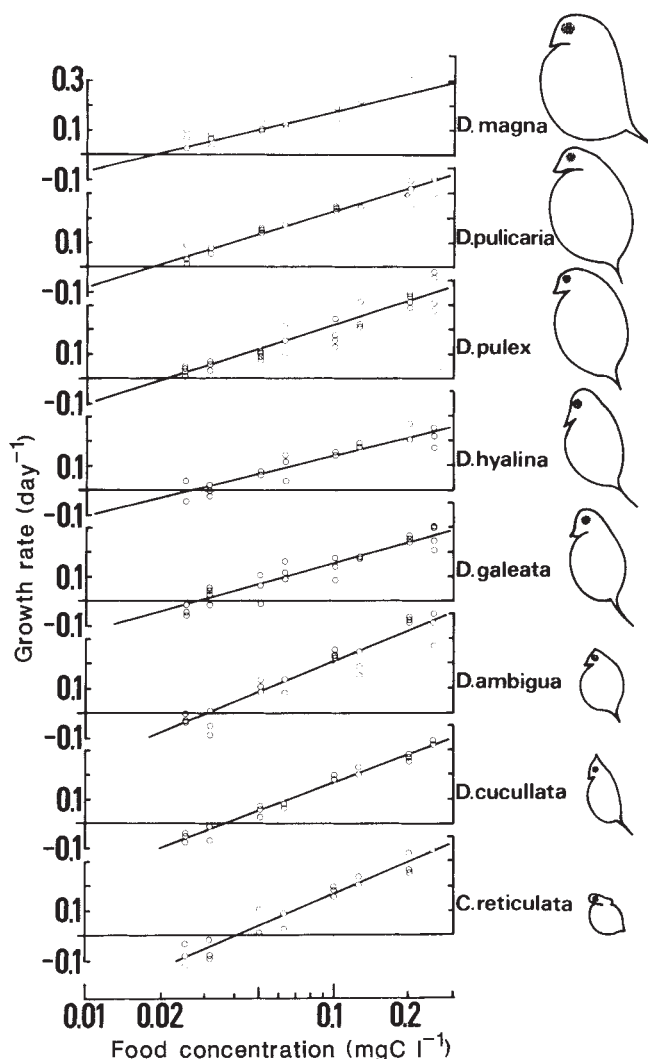
tion constant from the beginning to the end of experiment during the growth phase in which the animals were 2–6 days old. I identified the threshold food concentration of each species as the x-axis intercept of the growth rate (y-axis) on the natural log (ln) of the food concentration (x-axis) regression line, thereby giving the food concentration at zero growth rate.

I used the green algae *Scenedesmus acutus* MEYEN as a single food resource, thereby assuring that food particles, 8–

14 µm long, were well in the range of the particle size available to each of the eight species¹⁷. Each daphnid came from long-established clones at the Max-Planck-Institut für Limnologie in Plön, FRG, where they had been maintained in membrane-filtered (0.45 µm) lake water and fed *S. acutus* from a chemostat culture.

This approach, in which food levels are held constant in a flow-through experimental system (see Fig. 1), provided a

FIG. 1 Body growth rates of the eight daphnids at different food (*S. acutus*) concentrations. Values were determined as $g = (\ln C_t - \ln C_0)/t$, where C is the body carbon content of an individual at the beginning (C_0) or at the end (C_t) of the experimental time ($t = 4 \text{ days} \pm 0.2 \text{ day}$ resulting from the 10 h required to process all animals) when animals were 2 and 6 days old, respectively. In the highest food concentration (0.2 mg C l⁻¹ experiment 1; 0.25 mg C l⁻¹, experiments 2 and 3) 6-day-old animals were at the stage of releasing the first clutch of eggs into the brood chambers (Table 1). Daphnids were grown in a flow-through system²¹ which minimized the influence of grazing and sedimentation on concentration of algae. Chambers with 250 ml food suspension were maintained at 20 °C (± 0.1 °C). Food suspensions were pumped from stirred reservoirs by a multichannel peristaltic pump at the rate of 2 l per day per chamber. Fresh food suspensions were made every 24 h by diluting stock suspension with membrane-filtered lake water. Four final concentrations of 0.2, 0.1, 0.05 and 0.025 mg C l⁻¹ (experiment 1) or 0.25, 0.125, 0.062 and 0.031 mg C l⁻¹ (experiments 2 and 3) were selected according to previous estimates of food threshold levels for various *Daphnia* species^{12,13,23}. They were determined using calibration curves relating absorbance at 800 nm to organic carbon concentration. At the start of each experiment, 2-day-old daphnids were pipetted into chambers. They had been randomly selected from offspring released over 12 h 2 days previously. In experiment 1 (November 1988), each species was grown in separate chambers (three or four replicats) with 15–50 animals per chamber depending on body size and the related filtering rate¹⁶ (15 *D. magna*, 50 *D. cucullata* and *C. reticulata*). In experiments 2 and 3 (October and December, 1988, respectively), four species in various combinations were grown in the same chambers with 10 individuals of each species per chamber and always two larger and two smaller species together. At the end of each experiment, all 6 ± 0.2-day-old animals were collected, examined for body length and lipid reserves under a dissecting microscope, killed with formalin, rinsed in distilled water, dried, weighed and analysed for body carbon (Table 1). The slopes of all eight regressions are different from zero ($P < 0.001$, d.f. = 22, analysis of variance), but not from each other. Each growth rate estimate comes from a minimum of twenty 2-day-old or ten 6-day-old animals. The intercept of the regression line with the zero growth rate level can be underestimated for *D. magna* because of unexpectedly high values at the lowest food concentrations. The latter could have resulted from the ability of *D. magna* to use periphytic bacteria growing on the walls and bottom screens of chambers.



straightforward answer to the question of which species could grow at lower food concentrations. The answer, however, does not resolve the reasons for food limitation (low productivity of food species or high rate of their elimination by all coexisting grazers) or species-specific abilities to depress resources.

In all eight species, body growth rate was closely related to $\ln$ food concentration (Fig. 1). At the high food concentration of 0.20 to 0.25 mg C l^{-1} , growth rate (see Fig. 1, legend) was 0.3 per day in all species. This value is the same order of magnitude as the rate that I observed at the same temperature (20 °C) at nonlimiting *S. acutus* concentration of 0.5 mg C l^{-1} . At the low concentrations of 0.025 to 0.031 mg C l^{-1} , however, growth rate was clearly higher in large-bodied species than it was in small-bodied species, the latter often having negative growth rates. The lower body carbon values of 6-day-old versus 2-day-old *Ceriodaphnia reticulata*, *D. cucullata* and *D. ambigua* ($P < 0.02$, t test) were observed despite the fact that their bodies were significantly ($P < 0.01$, t test) longer (over 20%, 24% and 29%, respectively). The animals were, however, ghostly transparent and had no visible lipid reserves.

In Fig. 1 the intercepts of the regression lines with the zero growth rate level (assimilation equals respiration) indicate that food threshold concentrations are lower in larger species than in smaller species. The use of the independent variable (food concentration) to derive the estimated parameter (threshold food level) makes it statistically difficult to prove that the intercepts are actually different from each other. The lack of complete overlap in the 95% fiducial limits¹⁸ seen from the plot of the threshold values against specific body size (Fig. 2) makes it clear, however, that the intercepts of the two largest species (*D. magna* and *D. pulicaria*) are significantly different from the intercepts of the five smallest species, and that the intercepts of the three smallest species (*C. reticulata*, *D. cucullata* and *D. ambigua*) are significantly different from those of the three largest species.

Moreover, in Fig. 2 the slope of the regression of the threshold value on the species-specific body carbon of 6-day-old animals is also significant (analysis of variance, $P < 0.001$, degrees of freedom (d.f.) = 6), which makes it evident that the food threshold concentration increases with decreasing body size. The same was also observed when the threshold (intercept) values were plotted against the body carbon of 2-day-old animals and against dry body mass or body length of either 2- or 6-day-old animals (in all cases, $P < 0.001$, d.f. = 6).

The results described here are consistent with the cornerstone assumption of the size-efficiency hypothesis that the food assimilation should increase more rapidly with body size than does the respiration rate^{1,19}. This, however, seems to be correct only in a steady-state and single-resource situation which is neither likely to be encountered in nature nor frequently created

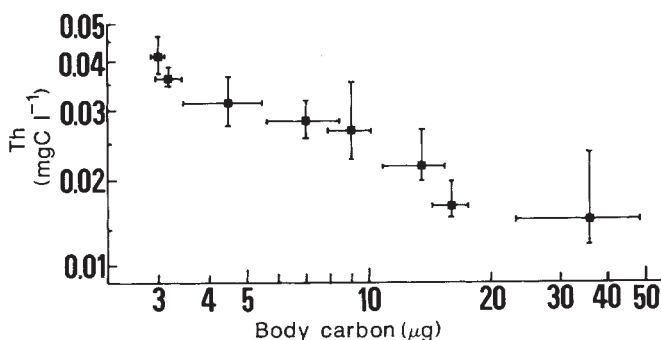


FIG. 2 Threshold food concentrations (Th) taken as the intercepts of the regression lines with the zero growth level from Fig. 1 with 95% fiducial limits, plotted against species-specific body size expressed as body carbon (means ± 1 s.d.) of 6-day-old animals grown at high food (*S. acutus*) levels (0.20 to 0.25 mg C l^{-1}). The slope of the regression ($\text{Th} = \ln(0.02)\text{C} - 2.85$) is different from zero ($P < 0.001$, d.f. = 6, analysis of variance).

in the laboratory where food concentration fluctuates (for example, in batch cultures where food is always depleted in the time between consecutive transfers of animals to fresh food suspensions²⁰). This could explain why small cladoceran species are superior competitors for limiting resources^{6,9,11,20}. Small-bodied species could be either more resistant to starvation and able to withstand longer periods of hunger at fluctuating food levels²¹, or they could be more resistant to interference from large phytoplankton²². The latter was recently shown for three *Daphnia* species of different body sizes²³. When grown on a readily available single-food resource, the largest of the three species, *D. pulicaria*, had the lowest food threshold concentration, and the smallest, *D. cucullata*, had the highest food threshold concentration. The intermediate-sized *D. hyalina* had a threshold between these two. The sequence, however, was reversed in the presence of high densities of a filamentous cyanobacterium, due to higher vulnerability of larger species to mechanical interference by filaments.

Although the assumption of the size-efficiency hypothesis about the competitive superiority of larger planktonic herbivores may be correct, the predictive value of the hypothesis is low because of the complex sources of nutrition, the possibility of food partitioning, the overlap among species in body size and fluctuating food levels in the field. It seems likely, however, that it is the ability of large *Daphnia* to depress the food resource level beyond the minimum necessary for smaller species to grow and reproduce, which makes it the only cladoceran resident in fishless temporary ponds (*D. magna*) or in large oligotrophic alpine lakes lacking fish (*D. pulicaria*)^{2,23}.

By estimating body growth between the second and the sixth day of animal life, I have demonstrated that the threshold food level decreases with an increase in body size. The estimate of body growth is an average value for a four-day history of fluctuations in assimilation and respiration. Further experiments should estimate simultaneously assimilation and respiration in animals from birth to maturity. This would clarify how species at a given age have a greater competitive ability. To show which species are competitively superior in nature, however, would require enclosure studies in lakes of different trophic status. Large inocula of all the *Daphnia* species considered here could be introduced simultaneously into the enclosure of each lake so that the relative performance of each large species and each small species, in the absence of predation, could be observed. □

Received 5 October 1989; accepted 8 January 1990.

- Brooks, J. L. & Dodson, S. I. *Science* **150**, 28–35 (1965).
- Gliwicz, M. Z. & Pijanowska, J. *Plankton Ecology: Succession in Plankton Communities* (ed. Sommer, U.) (Springer, Berlin, 1989).
- Kerfoot, W. C. (ed.) *Predation: Direct and Indirect Impacts on Aquatic Communities* (New England University Press, Hanover, New Hampshire, 1987).
- Zaret, T. M. *Predation and Freshwater Communities* (Yale University Press, 1980).
- Dodson, S. I. *Ecology* **55**, 605–613 (1974).
- Hanski, I. & Ranta, E. *J. Anim. Ecol.* **52**, 263–279 (1983).
- Lynch, M. *Ecology* **58**, 763–774 (1977).
- Lynch, M. *Limnol. Oceanogr.* **24**, 253–272 (1979).
- Lynch, M. in *Evolution and Ecology of Zooplankton Communities* (ed. Kerfoot, W. C.) (New England University Press, 1980).
- Romanovsky, Y. E. *Int. Rev. ges. Hydrobiol.* **69**, 613–632 (1984).
- Romanovsky, Y. E. *Arch. Hydrobiol.* **21**, 363–372 (1985).
- Tillmann, U. & Lampert, W. *J. Freshw. Ecol.* **2**, 311–323 (1984).
- Lampert, W. *Arch. Hydrobiol. (Suppl.)* **48**, 361–368 (1977).
- Rothhaupt, K. O. *Arch. Hydrobiol.* (in the press).
- DeMott, W. R. *Oecologia* **68**, 334–340 (1986).
- Lampert, W. in *Daphnia* (eds Peters, R. H. & deBernardi, R.) 143–192 (Mem. Ist. Ital. Idrobiol. Vol. 45, 1987).
- Geller, W. & Muller, H. *Oecologia* **49**, 316–321 (1981).
- Draper, N. R. & Smith, H. *Applied Regression Analysis* (Wiley, New York, 1981).
- Hall, D. J., Threlkeld, S. T., Burns, C. W. & Crowley, P. H. *A. Rev. Ecol. Syst.* **7**, 177–208 (1976).
- Tessier, A. J. & Goulden, C. E. *Limnol. Oceanogr.* **32**, 680–686 (1987).
- Stich, H. B. & Lampert, W. *Oecologia* **61**, 192–196 (1984).
- Gliwicz, Z. M. *Ecol. pol.* **25**, 179–225 (1977).
- Gliwicz, Z. M. & Lampert, W. *Ecology* (in the press).
- Krambeck, H. J., Lampert, W. & Brede, H. *Glas Instrument. Technik Fachz. Lab.* **25**, 1009–1012 (1981).

ACKNOWLEDGEMENTS. I thank Maren Volquardsen and Dietmar Lemke for their expert assistance in experimental work, and Mona Mort for comments on the manuscript.

Functional androdioecy in the flowering plant *Datisca glomerata*

Aaron Liston*, Loren H. Rieseberg & Thomas S. Elias

Rancho Santa Ana Botanic Garden, Graduate Program in Botany,
1500 North College Avenue, Claremont, California 91711-3101, USA

THE role of androdioecy (the presence of male and hermaphrodite individuals in a breeding population) in the evolution of dioecy has long been the subject of much interest and discussion¹⁻⁹. But no functionally androdioecious species has been previously documented² and recent studies have even raised doubt about whether the phenomenon exists at all³. Although many cases of androdioecy have been reported, most of these are based on morphological data alone and, when examined in detail, are actually found to be functionally dioecious¹⁰⁻¹². Here we describe functional androdioecy in the flowering plant *Datisca glomerata* (Presl.) Baill. (Datisceae). We suggest that the condition evolved from a dioecious precursor, and not from hermaphroditism as is commonly postulated for the evolution of androdioecy¹⁻⁹. Androdioecy in this case could be a transitional state in the breakdown of a dioecious breeding system towards hermaphroditism.

Datisca glomerata is distributed from northern California to northern Baja California, Mexico. Plants are long-lived tall (1–2.5 m) herbaceous perennials of riparian habitats. Flowering and fruiting occur throughout the summer months. The floral morphology of the species is relatively simple^{13,14}. Flowers are apetalous. Calyx lobes are short (1–1.5 mm) and inconspicuous. Male flowers have from six to twenty anthers. At anthesis, the pedicels of male flowers elongate, often reaching lengths of 2 cm or more. In hermaphrodite flowers the ovary is a three-sided capsule with three branched styles. They are protogynous and usually have three anthers. Nectaries are absent. The pendulous male flowers, large styles of the female and hermaphrodite flowers, unornamented pollen^{13,14} and the otherwise reduced flowers without nectar, are characteristic of an anemophilous pollination syndrome¹⁵. Likewise, the height and growth of the plant in an exposed riparian habitat are consistent with anemophily.

We verified the existence of functional androdioecy in *D. glomerata*. Hermaphrodite flowers in this species have fertile pollen (>95% reactive with an enzyme-specific pollen viability stain¹⁶), dehiscent anthers, and are self-compatible producing as many viable seeds on self-pollination as open pollinated flowers. Hermaphrodite flowers that we emasculated and bagged set no seed, indicating that non-pseudogamous apomixis does not occur. In addition, we examined 23 populations of *D. glomerata* for electrophoretically detectable genetic variation, revealing an inbreeding coefficient of $F = 0.617$ ($P < 0.001$)¹⁷ for six variable loci, indicating high inbreeding¹⁸. By contrast, we found no deviation from random mating ($F = -0.005$ ($P > 0.995$) for four variable loci) for three populations of the congeneric dioecious species *D. cannabina* L., native to south-western Asia¹⁷.

To confirm androdioecy in *Datisca*, it must be shown that individual plants are not sexually labile, that is, able to vary their sex according to their physiological condition or environment¹⁹⁻²¹. In a perennial plant with a prolonged flowering season such changes can be temporary and therefore difficult to detect. But flowers of *Datisca* do not abscise after anthesis, and their sex can be readily distinguished in the dry state. A mature inflorescence therefore contains a complete record of the sex of the flowers from the entire growing season. By observing all branches and inflorescences of 397 flowering plants at the end of the 1989 growing season, we found that plants remained

either hermaphrodite or male throughout the growing season (Table 1).

If physiological conditions determine the sex of plants, then we would have expected younger or smaller individuals to be male, and mature plants to be hermaphrodite²¹. But 24 plants grown from seed collected in two completely hermaphrodite populations flowered after two years and were all hermaphrodites. We observed three additional plants cultivated at the Rancho Santa Ana Botanic Garden over three growing seasons, and eight plants growing in the Big Tujunga Canyon population, and found no instances of year-to-year sex changes. Long-term observations of marked individuals in the experimental and natural populations are continuing.

We quantified the sex ratios in 10 populations and found that a 1:1 sex ratio, as would be expected in a functionally dioecious plant², does not exist between male and hermaphrodite individuals in *D. glomerata* (Table 1). In those populations in which we observed males, their frequency was always <25%. The finding of relatively few non-flowering individuals indicates that there is little potential for significant changes in these ratios due to recruitment.

One reason for the apparent rarity of androdioecious species is that male plants must have a high male fertility, at least double that of hermaphrodites, in order to be maintained by selection^{2,5,8}. An even greater advantage is required in partially self-fertilizing populations^{2,5,8} as found in *D. glomerata*, because the gain in fitness through having pollen is least when few ovules are available for outcrossing. Consistent with these stringent conditions for androdioecy, we found that male plants of *D. glomerata* have ~3.8 times as many fertile anthers as do hermaphrodites (male mean $x = 11.3 \pm 2.4$, $N = 97$; hermaphrodite $x = 3.0 \pm 0.6$, $N = 100$; both sexes have about equal numbers of flowers).

Despite the evidence for androdioecy in *D. glomerata*, this breeding system is very rare, and is not generally involved in the evolution of dioecy from hermaphroditism^{2,3}. In *Datisca*, it seems that androdioecy is in fact derived from dioecy. Evidence for this is found in the Datisceae, where the three species other than *D. glomerata* in the family are all dioecious. This includes *D. cannabina* and the presumed primitive members of the family, *Octomeles* and *Tetrameles*, forest trees of south-east Asia^{13,14}. A chloroplast DNA-based phylogeny of the family is now being reconstructed to test the hypothesis that androdioecy in the Datisceae is a derived condition.

It is interesting in this context that of 17 putatively (but undocumented) androdioecious taxa listed by Charlesworth², all but two have dioecious relatives. In two cases where nearly androdioecious species are thought to exist but no dioecy has

TABLE 1 No. individuals of each sex in populations of *D. glomerata*

Population	Hermaphrodite	Male	Non-flowering
Alder Creek	54	11	14
Baughman Spring	135	8	15
Big Tujunga Canyon	8	0	3
Cedar Springs Dam	30	1	2
Cloudburst Summit	3	0	0
Little Rock Dam	10	0	0
Mt Islip I	45	14	6
Mt Islip II	37	0	9
Tie Canyon I	26	2	0
Tie Canyon II	13	0	1

All populations are in the Angeles National Forest, Los Angeles County, California, USA, except for Cedar Springs Dam which is in the San Bernardino National Forest, San Bernardino County, California, USA. Populations were surveyed at the end of the flowering season to determine if sex changes had occurred over the course of the season as discussed in the text. The frequency of hermaphrodites and males in all populations shows a highly significant deviation from the 1:1 ratio expected with dioecy ($P < 0.001$). Mt, Mount.

* Present address: Department of Genetics, University of California, Davis, California 95616, USA.

been reported, heterostyly is present²²⁻²⁴. Heterostyly, like dioecy, represents an extreme form of outbreeding, and sub-androdioecy seems to be the outcome of a breakdown of the heterostylous condition²²⁻²⁴. We suggest that functional androdioecy in *Datisca* is the result of an analogous breakdown of the dioecious breeding system. The breakdown of dioecy could have resulted from pollination limitations due to low population densities (D. Charlesworth, personal communication).

In *D. glomerata*, female plants have already been replaced by self-compatible hermaphrodites. In this situation male individuals are at a strong reproductive disadvantage. Perhaps the large number of anthers and high pollen production of the males relative to the hermaphrodites has allowed the androdioecious condition to persist. Nevertheless, four of the populations of *D. glomerata* that we examined for sex ratios had already lost all male individuals (Table 1), and it seems likely that this process will continue, leading to a totally hermaphrodite breeding system. As such, androdioecy in *Datisca* probably represents a rare and transient state in the breakdown of dioecy. □

Received 25 July; accepted 14 December 1989.

1. Darwin, C. *The Different Forms of Flowers on Plants of the Same Species* (Murray, London, 1877).
2. Charlesworth, D. *Biol. J. Linn. Soc. Lond.* **23**, 333-348 (1984).
3. Anderson, G. J. & Symon, D. E. *Evolution* **43**, 204-219 (1989).
4. Bawa, K. S. A. *Rev. Ecol. Syst.* **11**, 15-39 (1980).
5. Lloyd, D. G. *Genetica* **45**, 325-339 (1975).
6. Ross, M. D. & Weir, B. S. *Evolution* **30**, 425-441 (1976).
7. Ross, M. D. *Am. Nat.* **119**, 297-318 (1982).
8. Charlesworth, B. & Charlesworth, D. *Am. Nat.* **112**, 975-997 (1978).
9. Westergaard, M. *Adv. Genet.* **9**, 217-281 (1958).
10. Schmid, R. *Bot. Jahrb. Syst.* **100**, 149-195 (1978).
11. Anderson, G. J. *Nature* **282**, 836-838 (1979).
12. Appanah, S. *Biol. J. Linn. Soc. Lond.* **18**, 11-34 (1982).
13. Davidson, C. thesis, Claremont Grad. Sch., Claremont, California (1972).
14. Davidson, C. *Aliso* **8**, 49-110 (1973).
15. Richards, A. J. *Plant Breeding Systems* (George Allen & Unwin, London, 1986).
16. Chandler, J. M., Jan, C. & Beard, B. H. *Syst. Bot.* **11**, 353-371 (1986).
17. Liston, A., Rieseberg, L. H. & Elias, T. S. *Aliso* **12**, 525-542 (1989).
18. Wright, S. *Evolution* **19**, 395-420 (1965).
19. Lloyd, D. G. & Bawa, K. S. *Evol. Biol.* **17**, 255-338 (1984).
20. Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
21. Schlessman, M. A. *Am. J. Bot.* **71**, 66 (abstr.) (1984).
22. Ornduff, R. *Am. J. Bot.* **51**, 307-314 (1964).
23. Ornduff, R. *Evolution* **26**, 52-65 (1972).
24. Ornduff, R. *Ann. Mo. Bot. Gdn* **70**, 146-148 (1983).

ACKNOWLEDGEMENTS. This work was supported by the Rancho Santa Ana Botanic Garden and a Sigma Xi award to A.L. Some of this study was carried out as part of the Bilateral Agreement on Environmental Protection between the United States and the Soviet Union. The help of the International Affairs Staff of the Fish and Wildlife Service is gratefully acknowledged.

Programmed death of autoreactive thymocytes

H. Robson MacDonald & Rosemary K. Lees

Ludwig Institute for Cancer Research, 1066 Epalinges, Switzerland

T LYMPHOCYTES bearing high-affinity T-cell receptors (TCR) for self-antigens are clonally deleted during thymus development¹. Several recent studies^{1,4-10} have identified variable domains of the β -chain of the TCR that are specifically deleted *in vivo* in mouse strains that express major histocompatibility complex class II molecules in addition to poorly defined self-antigens, including those encoded by the *Mls-1^a* and *Mls-2^a* loci. Deletion of autoreactive cells in these systems occurs in the thymus, and antibody blocking experiments *in vivo* have implicated the phenotypically immature CD4⁺CD8⁺ 'cortical' subset as the target population for clonal deletion^{11,12}. Similarly, studies with transgenic mice bearing autoreactive TCR have provided independent evidence that clonal deletion occurs at the CD4⁺CD8⁺ stage of development¹³. But none of these studies directly identified dying autoreactive cells, and the circumstances leading to deletion remain unclear. Here we report that neonatal thymus contains a significant popula-

tion of phenotypically mature CD4⁺CD8⁻ cells bearing autoreactive TCR. When placed in short-term culture, a large proportion (60%) of these autoreactive cells die selectively. Furthermore, their death can be prevented by inhibitors of macromolecule (RNA and protein) synthesis, as is the case for glucocorticoid-induced death of thymocytes^{2,3}. These data indicate that physiological clonal deletion of autoreactive cells involves 'programmed' cell death, and that it can occur in cells with a mature (CD4⁺CD8⁻) surface phenotype.

While investigating the ontogeny of expression of autoreactive β -chain of TCR (TCR V β), we have previously observed¹⁴ that V β 6⁺ cells with a mature (CD4⁺CD8⁻) phenotype could be transiently detected in the early postnatal thymus of BALB.D2.Mls^a (Mls-1^a) mice, although such cells were virtually absent from the thymus and periphery of adult animals. Further analysis of these CD4⁺ V β 6⁺ thymocytes revealed that they expressed approximately twofold fewer TCR than age-matched BALB/c (Mls-1^b) control mice, indicating that they could be undergoing receptor down-regulation as a consequence of interaction with a putative self-antigen¹⁴. This interpretation is greatly strengthened by the data presented in Table 1. When we placed CD8⁻ thymocytes isolated from 4-day-old BALB.D2.Mls^a mice *in vitro* at 37 °C in the absence of any deliberate stimulus, a reproducible fraction (~60%) of the CD4⁺ V β 6⁺ cells died within 20 h. We established by kinetic studies that death was detectable after 4 h, and we saw no increase in the fraction of dying V β 6⁺ cells by prolonging the culture period from 20 h to 48 h (Fig. 1). Death of CD4⁺ V β 6⁺ thymocytes was selective, because few ($\leq 10\%$) CD4⁺ cells died during this culture period and the levels of CD4⁺ cells expressing another TCR β -chain (V β 8) were unchanged (Table 1). Furthermore, CD4⁺ V β 6⁺ thymocytes from neonatal or adult BALB/c (Mls-1^b) mice did not die under comparable experimental conditions (Table 1).

Cell death in a variety of developmental systems involves chromatin condensation, membrane blebbing and fragmentation of DNA after the activation of a calcium-dependent endogenous endonuclease. These events have been collectively referred to as apoptosis¹⁵ and are believed to represent a programmed (as

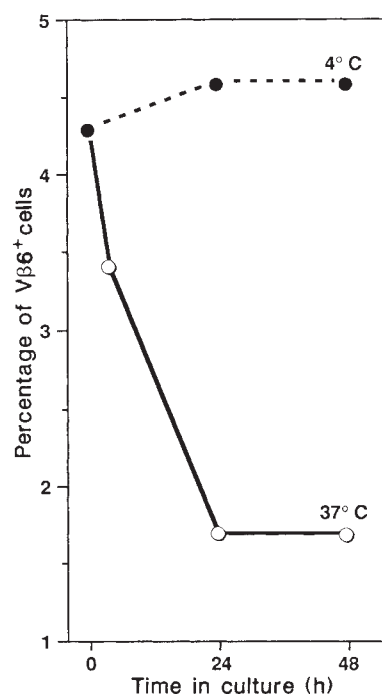


FIG. 1 Kinetics of death of autoreactive thymocytes *in vitro*. CD8-depleted thymocytes from 4-day-old BALB.D2.Mls^a mice were cultured at 4 °C or 37 °C. At various times, recovered viable cells were assessed for expression of V β 6 (see Table 1). No significant decrease in V β 8⁺ cells was observed at 37 °C at any time (Table 1, and data not shown).

TABLE 1 Selective death *in vitro* of CD4⁺ Vβ6⁺ thymocytes from neonatal BALB.D2.Mls^a mice

Thymocytes cultured	Percentage of Vβ6 ⁺ cells			Percentage of Vβ8 ⁺ cells		
	4 °C	37 °C	% reduction	4 °C	37 °C	% reduction
BALB.D2.Mls ^a (d4)	4.3	1.7	61	13.3	10.4	22
	3.4	1.0	71	11.6	11.9	-3
	3.3	1.5	55	13.4	11.6	13
	4.9	1.0	80	11.9	12.3	-3
	4.6	1.7	63	13.5	12.3	9
	4.2	1.9	55	11.3	10.2	10
	3.7	1.5	60	12.6	12.1	4
	4.3	2.3	47	10.8	9.7	10
	4.1 ± 0.6*	1.6 ± 0.4	62 ± 10	12.3 ± 1.0	11.3 ± 1.0	8 ± 8
BALB/c (d4)	9.4	10.6	-13	11.2	12.6	-13
	10.1	10.5	-4	12.5	12.3	2
	10.5	9.6	9	12.5	10.9	13
	10.0 ± 0.5	10.2 ± 0.5	-2 ± 11	12.1 ± 0.7	11.9 ± 0.9	1 ± 13
BALB/c (adult)	9.6	10.1	-5	19.9	17.2	14

CD8⁻ thymocytes from congenic BALB/c (Mls-1^b) or BALB.D2.Mls^a (Mls-1^a)²⁵ mice were cultured for 20 h at 4 °C or 37 °C. Viable recovered cells were stained with monoclonal antibodies directed against Vβ6 or Vβ8. Each row represents a separate experiment. Asterisk denotes mean ± s.d. Methods: Pooled thymi were depleted of CD8⁺ cells by treatment with rat IgM monoclonal antibody 3.168.1 in the presence of rabbit complement. Cells were cultured (1.5×10^6 ml⁻¹) in DMEM supplemented with 5% fetal bovine serum and 5×10^{-5} M 2-mercaptoethanol. Recovered cells were stained with rat monoclonal antibodies 44-22-1 (anti-Vβ6)²⁶, KJ16 (anti-Vβ8.1/8.2)²⁷ or H129.19 (anti-CD4)²⁸ followed by fluoresceinated goat anti-rat immunoglobulin. Viable cells were gated by a combination of forward and 90° light scatter²⁹. The percentage of positive cells was calculated by subtracting background staining with the fluorescent conjugate alone. Results were normalized in each case to the proportion of CD4⁺ cells in the CD8⁻ population, because this fraction increased after overnight culture at 37 °C (86 ± 6% at 37 °C versus 60 ± 4% at 4 °C, $n=8$). No detectable Vβ6⁺ cells were found among the remaining CD4⁻CD8⁻ population as assessed by two-colour fluorescence¹⁴.

opposed to accidental) mechanism of cell death normally occurring during development. Clonal deletion of autoreactive cells in the thymus is an obvious candidate for programmed cell death; indeed, circumstantial evidence favouring this model has been provided by the recent demonstration that apoptosis of thymocytes can be induced by administration of anti-CD3 monoclonal antibodies both *in vivo*¹⁶ and in thymus organ cultures¹⁷. But because anti-CD3 monoclonal antibodies are also potent inducers of apoptosis in target cells damaged by cytotoxic lymphocytes¹⁸, the interpretation of these studies remains equivocal.

To investigate further the mechanism of death of CD4⁺ Vβ6⁺ autoreactive thymocytes *in vitro*, we cultured these cells in the presence of inhibitors of macromolecular (RNA and protein) synthesis. Both actinomycin D and cycloheximide prevented the death of CD4⁺ Vβ6⁺ thymocytes when added at appropriate (non-toxic) concentrations (Fig. 2a). It is of interest that the same drugs inhibit glucocorticoid-induced death (and associated DNA fragmentation) in thymocytes, where they apparently block the synthesis of a so-far unidentified protein required for endonuclease activation^{2,3}.

Increased cytosolic Ca²⁺ concentration has also been implicated in endonuclease activation during glucocorticoid-mediated thymocyte death^{3,19}. In fact, DNA fragmentation (in total unseparated thymocytes) can be induced by calcium ionophores²⁰. Ionomycin did not further potentiate the death of autoreactive CD4⁺ Vβ6⁺ thymocytes in overnight cultures at 37 °C (Fig. 2b), although the dose of ionomycin used (0.5 μM) resulted in considerable death of CD4⁺CD8⁺ thymocytes in parallel cultures (data not shown). Such a result implies that contact with self-antigens induces sufficient calcium flux to kill autoreactive CD4⁺ Vβ6⁺ cells.

Taken together, these data provide strong evidence that autoreactive thymocytes undergo programmed cell death when they encounter physiological self-antigens during development. Also, they indicate that the dying autoreactive population includes at least some cells with a mature (CD4⁺CD8⁻) surface phenotype. But we emphasize that these results do not exclude the possibility that programmed cell death occurs at other (more immature) stages of T-cell development, nor do they imply that all T cells with a mature phenotype are susceptible to apoptosis. In this regard it is known that the primary target population *in vivo* of both glucocorticoids²¹ and anti-TCR antibodies¹⁷ is the

CD4⁺CD8⁺ thymocyte, and CD4⁺CD8⁺ cells have been implicated in clonal deletion in both normal^{11,12} and transgenic¹³ mice. Indeed, it is possible that the dying CD4⁺CD8⁻ Vβ6⁺ cells observed in the study described here are a relatively immature subset representing recent progeny of precursors that received an irreversible death signal at the CD4⁺CD8⁺ stage. Further characterization of the CD4⁺CD8⁻ Vβ6⁺ subset in neonatal BALB.D2.Mls^a thymus, particularly with respect to other markers that indicate their developmental status, will be required to address this issue.

Finally, the relationship of these findings to post-thymic (peripheral) tolerance remains to be explored. Recent studies from several groups^{22,23} are consistent with the concept that clonal anergy (rather than deletion) can account for functional tolerance of mature peripheral T lymphocytes bearing potentially autoreactive TCR. Also, in certain mouse strains that are

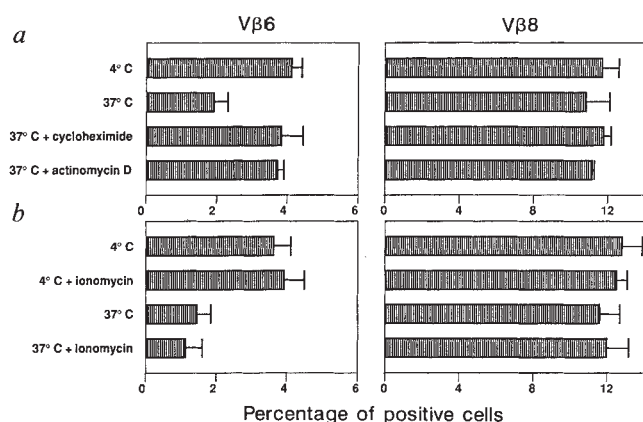


FIG. 2 Effect of inhibitors of macromolecular synthesis (a) or calcium ionophore (b) on death of autoreactive thymocytes *in vitro*. Values are mean ± s.d. CD8-depleted thymocytes from 4-day-old BALB.D2.Mls^a mice were cultured for 20 h at 4 °C or 37 °C. Cycloheximide (50 μg ml⁻¹), actinomycin D (5 μg ml⁻¹) or ionomycin (0.25 μg ml⁻¹) were added as indicated. Recovered viable cells were assessed for expression of Vβ6 and Vβ8 as described in Table 1. Data are from three independent experiments. Neither cycloheximide nor actinomycin D affected the viability of Vβ6⁺ or Vβ8⁺ cells maintained at 4 °C (data not shown).

prone to autoimmunity, expansion of autoreactive TCR V β domains occurs after early postnatal thymectomy²⁴. In the model system described here, CD4⁺ V β 6⁺ cells decrease with age in peripheral lymphoid tissues of neonatal BALB.D2.Mls^a mice¹⁴, and early postnatal thymectomy (performed after 3 or 6 days) does not result in a significantly increased proportion of CD4⁺ V β 6⁺ cells in adult animals (data not shown). The data described here therefore indicate that the expansion of potentially autoreactive T cells seen in autoimmune strains after neonatal thymectomy could depend on additional genetic factors (perhaps controlling the effective presentation of autoantigens). In strains where such clonal expansion does not occur (such as BALB.D2.Mls^a), it is nevertheless not clear whether the small proportion of residual CD4⁺ V β 6⁺ cells in the periphery represents an anergic population or alternatively a small subset of V β 6⁺ cells that has escaped clonal deletion because of lack of reactivity to Mls-1^a antigens. □

Received 19 October; accepted 19 December 1989.

- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 263–271 (1987).
- Wyllie, A. H., Morris, R. G., Smith, A. L. & Dunlop, D. *J. Path.* **142**, 67–77 (1984).
- Cohen, J. J. & Duke, R. C. *J. Immun.* **132**, 38–42 (1984).
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. *Nature* **332**, 35–40 (1988).
- MacDonald, H. R. *et al. Nature* **332**, 40–45 (1988).
- Pullen, A. M., Marrack, P. & Kappler, J. W. *Nature* **335**, 796–801 (1988).
- Fry, A. M. & Matis, L. A. *Nature* **335**, 830–832 (1988).
- Tomonari, K. & Lovering, E. *Immunogenetics* **28**, 445–451 (1988).
- Bill, J., Kanagawa, O., Woodland, D. L. & Palmer, E. *J. exp. Med.* **169**, 1405–1419 (1989).
- Kanagawa, O., Palmer, E. & Bill, J. *Cell Immun.* **119**, 412–426 (1989).
- Fowlkes, B. J., Schwartz, R. H. & Pardoll, D. M. *Nature* **334**, 620–623 (1988).
- MacDonald, H. R., Hengartner, H. & Pedrazzini, T. *Nature* **335**, 174–176 (1988).
- Kisielow, P., Bluthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742–746 (1988).
- Schneider, R. *et al. J. exp. Med.* **169**, 2149–2158 (1989).
- Wyllie, A. H., Kerr, J. F. R. & Currie, A. R. *Int. Rev. Cytol.* **68**, 251–305 (1980).
- Shi, Y., Sahai, B. M. & Green, D. R. *Nature* **339**, 625–626 (1989).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181–184 (1989).
- Russell, J. H. *Immunol. Rev.* **72**, 97–118 (1983).
- McConkey, D. J. *et al. Archs Biochem. Biophys.* **269**, 365–370 (1989).
- McConkey, D. J., Hartzell, P., Nicotera, P. & Orrenius, S. *FASEB J.* **3**, 1843–1849 (1989).
- Ceredig, R., Dyalinas, D. P., Fitch, F. W. & MacDonald, H. R. *J. exp. Med.* **158**, 1654–1671 (1983).
- Qin, S., Cobbold, S., Benjamin, R. & Waldmann, H. *J. exp. Med.* **169**, 779–794 (1989).
- Rammensee, H.-G., Kroschewski, R. & Frangoulis, B. *Nature* **339**, 541–544 (1989).
- Smith, H., Chen, I.-M., Kubo, R. & Tung, K. S. K. *Science* **245**, 749–752 (1989).
- Festenstein, H. & Berumen, L. *Transplantation* **37**, 322–324 (1984).
- Payne, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7695–7698 (1988).
- Haskins, K. *et al. J. exp. Med.* **160**, 452–471 (1984).
- Pierres, A. *et al. J. Immun.* **132**, 2775–2782 (1984).
- MacDonald, H. R. & Zaech, P. *Cytometry* **3**, 55–58 (1982).

ACKNOWLEDGEMENTS. We thank P. Zaech and C. Knabenhans for flow cytometry and A. Zoppi for preparation of the manuscript. We also acknowledge the contribution of R. Schneider and H. Hengartner (Institute for Pathology, University Hospital, Zurich) to the initial stages of this work.

A dimension reduction framework for understanding cortical maps

Richard Durbin* & Graeme Mitchison†

* Department of Psychology, Stanford University, California 94305, USA

† Physiological Laboratory, Downing Street, Cambridge CB2 3EG, UK, and The Research Centre, King's College, Cambridge CB2 1ST, UK

WE argue that cortical maps, such as those for ocular dominance, orientation and retinotopic position in primary visual cortex¹, can be understood in terms of dimension-reducing mappings from many-dimensional parameter spaces to the surface of the cortex. The goal of these mappings is to preserve as far as possible neighbourhood relations in parameter space so that local computations in parameter space can be performed locally in the cortex. We have found that, in a simple case², certain self-organizing models^{3,4} generate maps that are near-optimally local, in the sense that they come close to minimizing the neuronal wiring required for local operations. When these self-organizing models are applied to the task of simultaneously mapping retinotopic position and orientation, they produce maps with orientation vortices resem-

bling those produced in primary visual cortex⁵. This approach also yields a new prediction, which is that the mapping of position in visual cortex will be distorted in the orientation fracture zones⁵.

In addition to the retinotopic map in primary visual cortex (V1), orientation selectivity and ocular dominance are mapped in patterns of stripes or patches¹, as has been strikingly demonstrated by recent work⁵ using voltage-sensitive dyes (Fig. 2a). A characteristic feature of all these mappings is that they are smooth, that is, neighbouring cells on the cortex have similar response properties. Because receptive field properties are columnar, that is, are similar in a penetration normal to the cortical surface⁶, the cortex can be considered as a two-dimensional (2-D) sheet on which response properties are mapped. We assume that there are cells in V1 that are responsive to stimuli of each orientation in each eye at each position in visual space. We should therefore envisage a parameter space of response properties whose dimensions are orientation, ocular dominance and retinotopic position, which is mapped onto the cortical sheet. A similar organization, in which several response properties are all smoothly mapped into one area, is seen in other cortical areas^{7,8}.

The smoothness of cortical maps tells us that neighbouring points on the cortex correspond to nearby points in parameter space. But we propose here, following the arguments of Cowey⁹, that the main principle of cortical mapping operates the other way around, that is, neighbouring points in parameter space should map close together on the cortex. The underlying assumption is that most operations performed in the cortex are local, in the sense that they involve neighbouring parameter values. The construction of inhibitory surrounds and the sharpening of orientation tuning¹⁰ are examples of this. Such operations require connections between neurons in the cortex that represent similar parameter values. Here we use the length of cortical wiring needed to connect these neurons as a measure of how locally the computation is performed on the cortex. Minimizing wire length will ensure that cells representing similar parameter values lie close together. This approach is not restricted to V1. In general, it can be argued that those properties that are mapped in a given cortical area are precisely those that will be involved in local computational interactions, such as the sharpening of receptive field characteristics or extraction of new properties.

The parameter space will generally be of higher dimension than the 2-D cortical sheet. Thus in V1 there are at least four dimensions: two of retinotopic space, one of orientation and one of ocular dominance. Because of the reduction in dimension in mapping from parameter space to cortex, the requirement that all neighbouring points in parameter space should be neighbours on the cortex cannot be met. What is the best that can be achieved?

We have previously studied this problem² in a highly simplified setting where a 2-D discrete parameter space ($N \times N$) is mapped onto a 1-D 'cortex'. The way we formulated this problem and typical results are shown in Fig. 1. Our main conclusion is that optimal dimension reduction mappings show many of the features of cortical maps, but only if the mapping cost is calculated using a nonlinear function of wire length that favours many short connections while allowing a few long ones. The basic properties that the simplified model predicts are that the maps on the cortex are smooth, and that both parameters show jitter to accommodate mapping of the other parameter.

Unfortunately, it is computationally unfeasible to extend this kind of minimization procedure to higher-dimensional mappings to obtain optimal mappings on a 2-D model cortex. But this might not have been profitable anyway; we cannot claim to have an accurate measure of biological cost, nor is it likely that an explicit cost minimization process occurs during cortical development. Our strategy was therefore to examine self-organizing algorithms as models for the development of cortical maps. These have a much better claim to biological plausibility, and we show below that the maps they produce can achieve a

low cost by our wiring criterion, and that they can also generate maps of orientation like those seen in V1.

The basic assumption of these models, which has much support from sensory physiology¹¹⁻¹⁴, is that neurons develop their receptive-field characteristics under the influence of incoming neural activity. The rules for adjusting the response properties of cortical units are (1) that there are competitive interactions between units that correspond to any given input, after which those units that fire most strongly become more responsive to the input, and (2) that units also strengthen their responses to inputs that their neighbours respond to. The first rule leads to a competitive sharing-out of the domain of inputs among units, while the second rule imposes a continuity constraint leading to the development of a coherent map. Such algorithms have been able to account for many of the phenomena of plasticity of neural maps. The details of the algorithm we use are given in the legend to Fig. 2.

To what extent does this self-organizing procedure minimize wire length? If we apply our algorithm to the simplified problem described above (2-D parameter space to 1-D cortex), we obtain

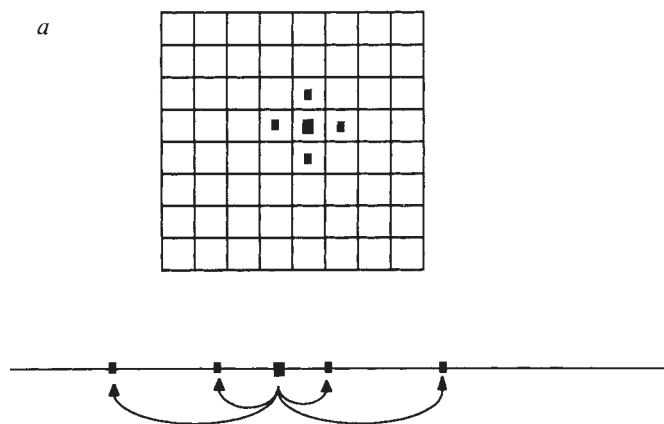
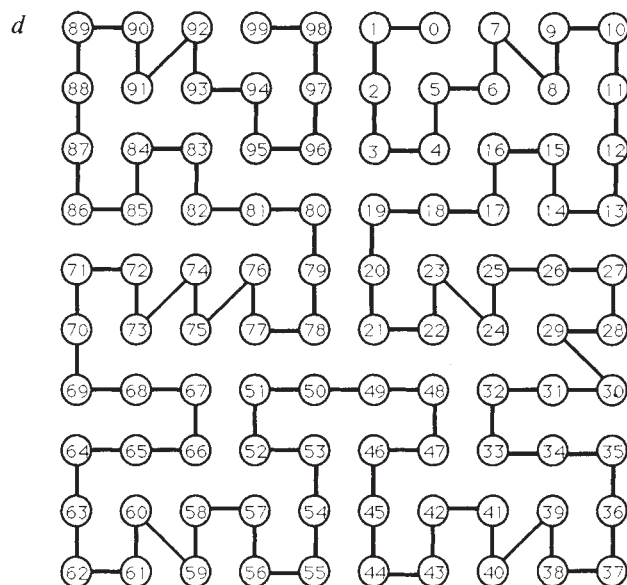
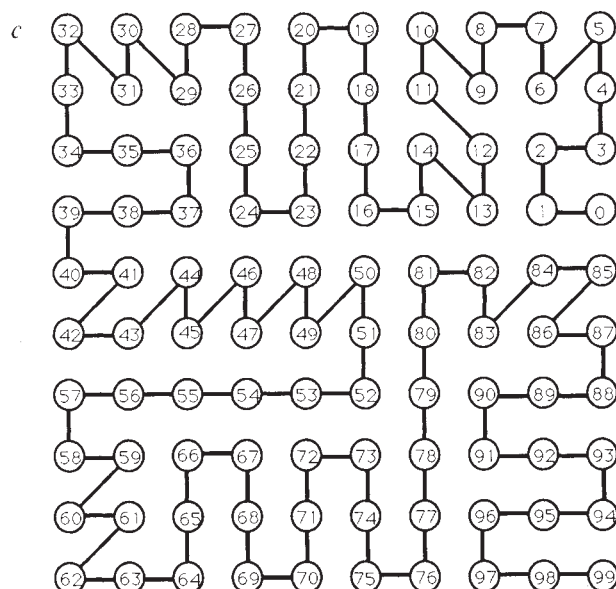
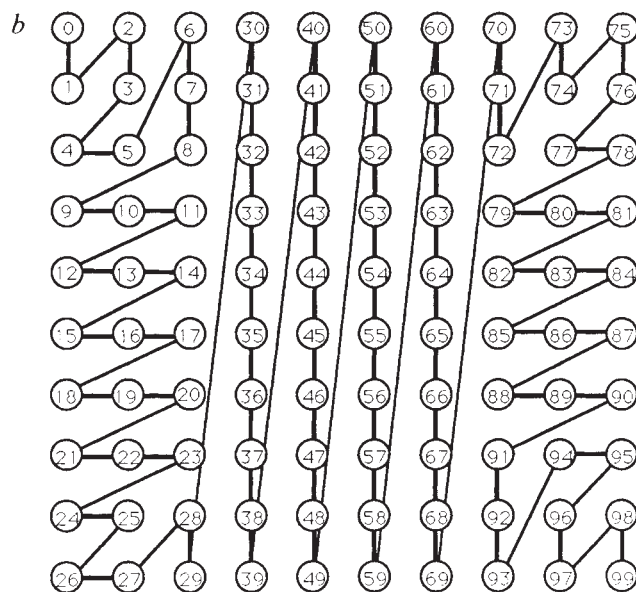


FIG. 1 *a*, Simplified model for dimension reduction. Points in the $N \times N$ grid are mapped to the line cortex by the function f . A point (large dot) and its four neighbours (small dots) are shown in the grid, and their images under f are shown on the line below. The 'wire length' for the mapping is the sum of the absolute values of differences between images of adjacent points, raised to the power p (L_p metric). Thus the link between $f(i, j)$ and $f(i, j+1)$ contributes $|f(i, j) - f(i, j+1)|^p$ to the cost. If $p > 1$, then this penalizes long-distance costs heavily, whereas if $p < 1$, it places an emphasis on mapping as many neighbours as close as possible, while allowing a few longer connections. The dependence of the optimal mapping on p is discussed in detail in ref. 2. *b*, Mapping found by applying our cost-minimizing algorithm to a 10×10 array with $p = 1$. The algorithm works by a structured search process involving generation and selection of random variants (4×10^7 operations in total). The mapping is depicted by showing the value of f (the integer within each circle) at each grid location. To allow the sequence of numbers to be followed more easily, lines have been drawn between locations with adjacent values. The mapping shown here can be proved to be optimal², with cost 914. The main characteristic of the mapping, shown clearly in the middle region, is that as the integers are run through (equivalent to moving along the cortex), the values of one parameter are rapidly cycled through with regular jumps, whereas the other advances one notch at a time. This is unlike cortical maps, which change more continuously, although with some random scatter. *c*, Mapping found by applying the cost-minimizing algorithm described above, minimizing with $p = \frac{1}{2}$. The organization is typical of mappings obtained when $p < 1$ (ref. 2). The cost is 335.8. This mapping has the fine-scale cortical properties that were lacking in *b*: smoothness (neighbouring numbers are nearby in the grid), and more equal treatment of the two parameters. *d*, Mapping found by applying the self-organizing algorithm (see legend to Fig. 2*b*). The cost ($p = \frac{1}{2}$) is 347.1. This is close to that for the optimized mapping shown in *c*, and many of the general features of the mapping in *c* are also seen here. For comparison, the cost of the mapping in *b* with $p = \frac{1}{2}$ is 352.3. Although for a 10×10 array the costs of *b* and *d* are close, the difference will increase with N , because the cost of *b* scales as $N^{2.5}$, whereas that for the more folded structures of *c* and *d* is expected to scale as $N^2 \log N$.



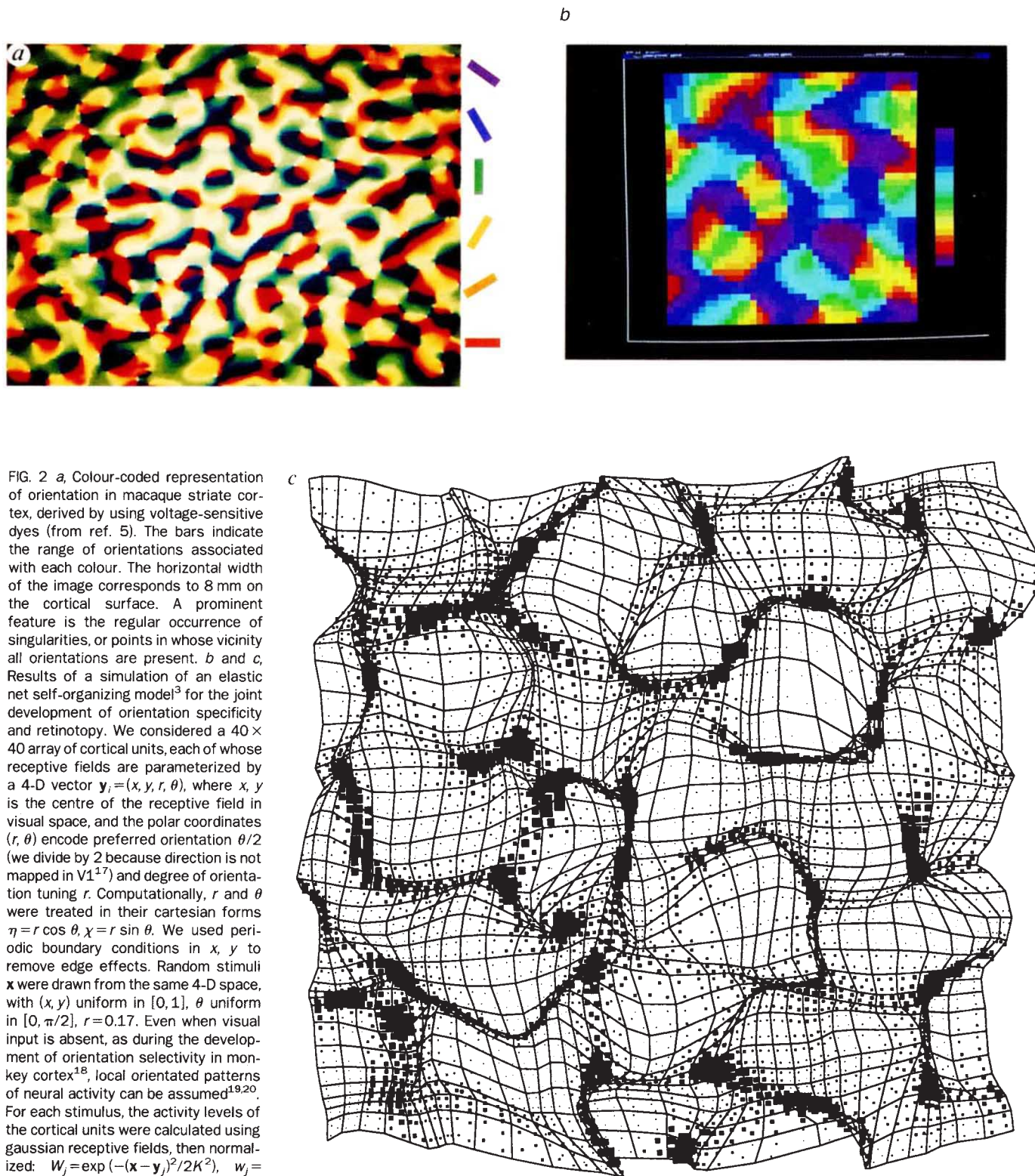


FIG. 2 *a*, Colour-coded representation of orientation in macaque striate cortex, derived by using voltage-sensitive dyes (from ref. 5). The bars indicate the range of orientations associated with each colour. The horizontal width of the image corresponds to 8 mm on the cortical surface. A prominent feature is the regular occurrence of singularities, or points in whose vicinity all orientations are present. *b* and *c*, Results of a simulation of an elastic net self-organizing model³ for the joint development of orientation specificity and retinotopy. We considered a 40×40 array of cortical units, each of whose receptive fields are parameterized by a 4-D vector $\mathbf{y}_i = (x, y, r, \theta)$, where x, y is the centre of the receptive field in visual space, and the polar coordinates (r, θ) encode preferred orientation $\theta/2$ (we divide by 2 because direction is not mapped in V1¹⁷) and degree of orientation tuning r . Computationally, r and θ were treated in their cartesian forms $\eta = r \cos \theta$, $\chi = r \sin \theta$. We used periodic boundary conditions in x, y to remove edge effects. Random stimuli $\mathbf{x}$ were drawn from the same 4-D space, with (x, y) uniform in $[0, 1]$, θ uniform in $[0, \pi/2]$, $r = 0.17$. Even when visual input is absent, as during the development of orientation selectivity in monkey cortex¹⁸, local orientated patterns of neural activity can be assumed^{19,20}. For each stimulus, the activity levels of the cortical units were calculated using gaussian receptive fields, then normalized: $W_j = \exp(-(\mathbf{x} - \mathbf{y}_j)^2/2K^2)$, $w_j = W_j/\sum W_k$. Receptive fields of all units j were then updated according to $\Delta \mathbf{y}_j = \alpha w_j(\mathbf{x} - \mathbf{y}_j) + \beta \sum_{k \in N} (y_k - y_j)$, where k in the sum ranges over the set N of direct neighbours to unit j in the cortical array. The first term in $\Delta \mathbf{y}_j$ is Hebbian. The second tends to make neighbours have similar receptive fields. Parameter values were $\alpha = 0.4$, $\beta = 0.0001$, $K = 0.06$. In fact, for reasons of computational efficiency, three sets of 216 random stimuli were used, and for each set a Gauss-Seidel procedure was used to rapidly find the stable configuration corresponding to repeated selection in the above procedure of the stimuli in the set. *b*, A colour-coded representation of the derived orientation map on the cortex. The magnification is about three

times higher than in Fig. 2*a*. The strip shows the colours corresponding to a 180-degree cycle of orientations. *c*, A map in visual space showing simultaneously the distortion of the retinotopic map and the rate of change of orientation. The distorted grid represents the cortical array of cells. Each grid intersection is at the receptive field centre of the corresponding cell. The small filled squares show the rate of change of orientation at the corresponding point on the cortical map. The diameters of the squares are proportional to the sum of differences of orientation between a cell and a horizontal and vertical neighbour.

maps (for example, Fig. 1d) that look very like the computed minimal cost mapping (Fig. 1c). The costs of these maps are indeed close to minimal (Fig. 1 legend). It is surprising that the self-organizing models achieve such a low cost, because they actually operate in the opposite direction to our minimal wiring problem. They explicitly map from cortex to parameter space, assigning parameter values to cortical points, and attempting to put nearby cortical points close in parameter space. The goal of the wiring problem is to map from parameter space to cortex, putting nearby points in parameter space as close as possible on the cortex. In fact, the problem that the self-organizing models explicitly solve, that of finding smooth mappings from cortex to parameter space, is underconstrained—there are many such smooth mappings. The particular solutions that the models find are also good solutions to the wiring problem, because the way the mapping develops gives it a folded structure, which has the incidental effect of allowing many short connections between adjacent parameter values. We suggest, therefore, that self-organizing procedures could have evolved as a way of obtaining a satisfactory solution to the wiring problem at much less expense than a more direct minimization would require.

To help justify this claim, and the dimension reduction approach in general, we show below how a dimension-reducing self-organizing model predicts some of the observed features of cortical maps. Here we simulate the joint representation of retinal position and orientation selectivity, ignoring, for simplicity, ocular dominance and other V1 parameters, such as colour. We consider a 2-D cortex and a 4-D parameter space in which retinal position (2-D), orientation and orientation tuning (2-D) are represented. The map generated is shown in Fig. 2b and c (details of the procedures used are given in the legend to Fig. 2).

As can be seen, the orientation map (Fig. 2b) is qualitatively not unlike that found experimentally (Fig. 2a). We note that the map contains singularities like those seen in the data of Blasdel and Salama⁵. Figure 2c shows how retinal position is mapped; the cortex is depicted here as a grid mapped onto retinal space. Superposed on this grid are squares whose sizes represent the rate of change of orientation. We note that rapid orientation change is not confined to the singularities, but extends out to form bridges between them. This is roughly comparable to the 'fractures' described by Blasdel and Salama⁵, although the rate of change is less in the stripes than at the singularities. Swindale has recently shown that his model for development of orientation selectivity¹⁶ also gives rise to these features (personal communication). It therefore does not seem to be necessary to postulate a qualitative difference between the fracture zones and the domains. They could be a consequence of the developmental mechanism that gives rise to the orientation map.

A novel consequence of our approach is that because retinal space is treated as one of the parameter space variables, on the same footing as orientation, there are distortions in the retinotopic map. The most rapid changes of orientation tend to occur where retinal position changes most slowly (at the folds of the grid in Fig. 2c). From this it can be predicted that in the fracture zones, position will change slowly with more backtracking or jitter, whereas in the areas of more constant orientation selectivity, it will vary more smoothly and more rapidly. It will be interesting to see if systematic fine-scale mapping of retinotopy together with orientation supports this prediction. More generally, our approach indicates that wherever several parameters that are jointly involved in local computation are mapped, the fine structure of their mapping may be correlated. A previously observed example of this type of interaction is that orientation vortices tend to be positioned along the middle of ocular dominance stripes⁵.

Our conclusion is that keeping connection lengths to a minimum could be an important goal of cortical map development, and that self-organizing algorithms could have evolved

as mechanisms that are able to generate very good, although not perfect, solutions to this problem. We have shown here an example of how developmental models based on such self-organizing algorithms can have both explanatory and predictive power. □

Received 2 October; accepted 19 December 1989.

- Hubel, D. H. & Wiesel, T. N. *Proc. R. Soc. B* **198**, 1–59 (1977).
- Mitchison, G. J. & Durbin, R. M. *SIAM J. alg. disc. Meth.* **7**, 571–582 (1986).
- Durbin, R. M. & Willshaw, D. *Nature* **326**, 689–691 (1987).
- Kohonen, T. *Self-organization and Associative Memory* (Springer, Berlin, 1984).
- Blasdel, G. G. & Salama, G. *Nature* **321**, 579–585 (1986).
- Mountcastle, V. B. *J. Neurophysiol.* **20**, 408–434 (1957).
- Kaas, J. H., Nelson, R. J., Sur, M. & Merzenich, M. M. in *The Organization of the Cerebral Cortex* (eds Schmitt, F. O., Worden, F. G., Adelman, G. & Denis, S. G.) 237–261 (MIT, Cambridge, Massachusetts, and London, 1981).
- Suga, N. & O'Neill, W. E. *Science* **206**, 351–353 (1979).
- Cowey, A. Q. *J. exp. Psychol.* **31**, 1–17 (1979).
- Hata, Y., Tsumoto, T., Sato, H., Hagihara, K. & Tamura, H. *Nature* **335**, 815–817 (1988).
- Hubel, D. J., Wiesel, T. N. & LeVay, S. *Phil. Trans. R. Soc. B278*, 377–409 (1977).
- Van der Loos, H. & Woolsey, T. A. *Science* **179**, 395–397 (1973).
- Merzenich, M. M. et al. *J. comp. Neurol.* **224**, 591–605 (1984).
- Stryker, M. P. & Harris, W. A. *J. Neurosci.* **6**, 2117–2133 (1986).
- Ritter, H. & Schulten, K. *Biol. Cybern.* **54**, 99–106 (1986).
- Swindale, N. V. *Proc. R. Soc. B215*, 211–230 (1982).
- Hubel, D. J. & Wiesel, T. N. *J. Physiol.* **195**, 215–243 (1968).
- Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **158**, 307–318 (1974).
- Linsker, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8390–8394 (1986).
- Miller, K. D., Keller, J. B. & Stryker, M. P. *Science* **245**, 605–615 (1989).

ACKNOWLEDGEMENTS. We thank Nick Swindale and Ken Miller for comments. R.M.D. is a Lucille P. Markey Visiting Fellow, and this work was supported in part by the Lucille P. Markey Charitable Trust.

Calcium/calmodulin-dependent protein kinase II increases glutamate and noradrenaline release from synaptosomes

Robert A. Nichols, Talvinder S. Sihra,
Andrew J. Czernik, Angus C. Nairn
& Paul Greengard

Laboratory of Molecular and Cellular Neuroscience,
The Rockefeller University, New York, New York 10021, USA

A VARIETY of evidence indicates that calcium-dependent protein phosphorylation modulates the release of neurotransmitter from nerve terminals^{1–3}. For instance, the injection of rat calcium/calmodulin-dependent protein kinase II (Ca²⁺/CaM-dependent PK II) into the preterminal digit of the squid giant synapse leads to an increase in the release of a so-far unidentified neurotransmitter induced by presynaptic depolarization^{3,4}. But until now, it has not been demonstrated that Ca²⁺/CaM-dependent PK II can also regulate neurotransmitter release in the vertebrate nervous system. Here we report that the introduction of Ca²⁺/CaM-dependent PK II, autoactivated^{5–8} by thiophosphorylation, into rat brain synaptosomes (isolated nerve terminals) increases the initial rate of induced release of two neurotransmitters, glutamate and noradrenaline. We also show that introduction of a selective peptidergic inhibitor of Ca²⁺/CaM-dependent PK II inhibits the initial rate of induced glutamate release. These results support the hypothesis^{3,9} that activation of Ca²⁺/CaM-dependent PK II in the nerve terminal removes a constraint on neurotransmitter release.

Because injected Ca²⁺/CaM-dependent PK II can enhance the release of neurotransmitter at the squid giant synapse³, we investigated whether a similar phenomenon is observed when the enzyme is introduced into nerve terminals from mammalian brain. To do this, we used a novel freeze-thaw technique that causes a transient permeabilization of rat brain synaptosomes while preserving their viability¹⁰. To maintain an increased

intrasyntosomal activity of Ca^{2+} /CaM-dependent PK II, we used enzyme preparations that we had activated by autophosphorylation (thereby generating a Ca^{2+} /CaM-independent form)⁵⁻⁸ before introducing them by the freeze-thaw technique. We autothiophosphorylated the enzyme using ATP- γ -S so that a phosphorylated form of the protein kinase that is resistant to phosphatase¹¹ was obtained. The activation of purified Ca^{2+} /CaM-dependent PK II obtained by autothiophosphoryla-

tion is comparable to that obtained by autophosphorylation (ref. 12; F. Gorelick & P.G., unpublished observations).

After introduction of autoactivated Ca^{2+} /CaM-dependent PK II into rat cortical synaptosomes, the initial rate of endogenous glutamate release induced by the Ca^{2+} ionophore ionomycin was increased to 152% of control (Fig. 1a). (Ionomycin has been shown to stimulate the release of glutamate, after a brief time lag, at rates comparable to those obtained with

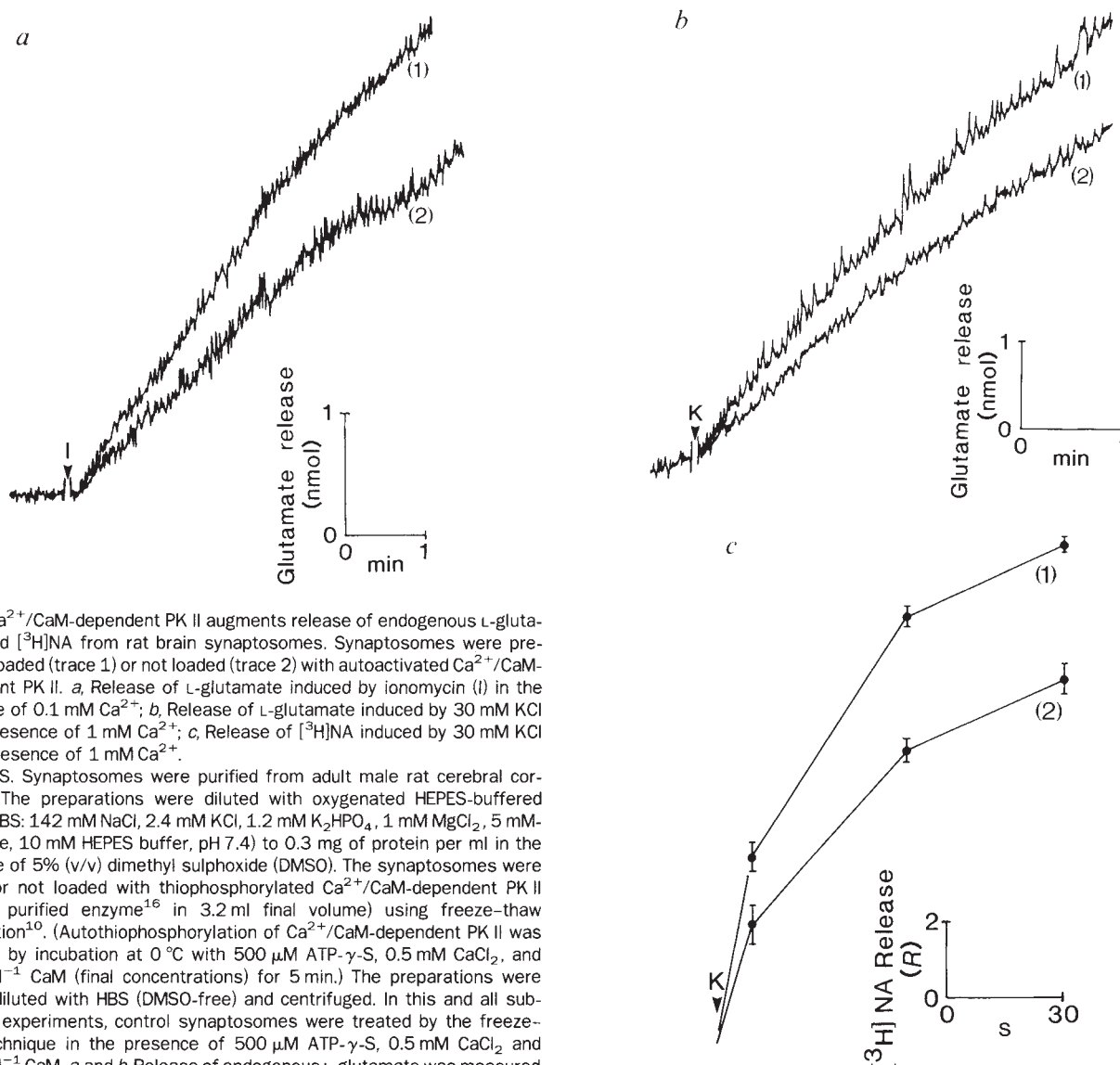


FIG. 1 Ca^{2+} /CaM-dependent PK II augments release of endogenous L-glutamate and $[^3\text{H}]\text{NA}$ from rat brain synaptosomes. Synaptosomes were previously loaded (trace 1) or not loaded (trace 2) with autoactivated Ca^{2+} /CaM-dependent PK II. *a*, Release of L-glutamate induced by ionomycin (I) in the presence of 0.1 mM Ca^{2+} ; *b*, Release of L-glutamate induced by 30 mM KCl in the presence of 1 mM Ca^{2+} ; *c*, Release of $[^3\text{H}]\text{NA}$ induced by 30 mM KCl in the presence of 1 mM Ca^{2+} .

METHODS. Synaptosomes were purified from adult male rat cerebral cortices²⁶. The preparations were diluted with oxygenated HEPES-buffered saline (HBS: 142 mM NaCl, 2.4 mM KCl, 1.2 mM K_2HPO_4 , 1 mM MgCl_2 , 5 mM D-glucose, 10 mM HEPES buffer, pH 7.4) to 0.3 mg of protein per ml in the presence of 5% (v/v) dimethyl sulphoxide (DMSO). The synaptosomes were loaded or not loaded with thiophosphorylated Ca^{2+} /CaM-dependent PK II (200 μg purified enzyme¹⁶ in 3.2 ml final volume) using freeze-thaw introduction¹⁰. (Autothiophosphorylation of Ca^{2+} /CaM-dependent PK II was achieved by incubation at 0 °C with 500 μM ATP- γ -S, 0.5 mM CaCl_2 , and 30 μg ml⁻¹ CaM (final concentrations) for 5 min.) The preparations were further diluted with HBS (DMSO-free) and centrifuged. In this and all subsequent experiments, control synaptosomes were treated by the freeze-thaw technique in the presence of 500 μM ATP- γ -S, 0.5 mM CaCl_2 and 30 μg ml⁻¹ CaM. *a* and *b*, Release of endogenous L-glutamate was measured using a continuous fluorometric assay²⁷. The synaptosomes (1 mg protein) were resuspended at 0.67 mg ml⁻¹ in 1.5 ml HBS and transferred to a stirred quartz cuvette maintained at 37 °C within a Perkin-Elmer LS-5B luminescence spectrometer. After 1 min, NADP⁺ (1 mM final concentration) was added and the preparation incubated for 2 min. L-Glutamate dehydrogenase (Sigma; EC 1.4.1.3, 50 U) was added and the preparation incubated for a further 5 min. Ionomycin (I; 5 μM) or 30 mM KCl (K) was added when indicated (arrowhead). Released L-glutamate is converted by L-glutamate dehydrogenase to α -ketoglutarate with the concomitant reduction of NADP⁺ to NADPH, the latter producing an increase in fluorescence (excitation, 330 nm; emission, 460 nm). The changes in fluorescence due to NADPH formation were calibrated (vertical bars) against additions of known L-glutamate standards. As previously shown for release of radiolabelled neurotransmitter¹⁰, the fractional release *R* (defined here as the ratio of the initial rate of L-glutamate release, expressed as nmol per min per mg protein, to the total L-glutamate pool, expressed as nmol per mg) of endogenous L-glutamate from synaptosomes was preserved after freeze-thaw treatment in the presence of DMSO ($R=0.055$) when compared with synaptosomes that were either untreated ($R=0.058$) or treated with DMSO

but not treated by the freeze-thaw technique ($R=0.068$). (In all experiments, the total L-glutamate level was determined at the end of each run after addition of Triton X-100 to release the total L-glutamate pool²⁸.) *c*, K⁺-induced Ca^{2+} -dependent release of $[^3\text{H}]\text{NA}$ was measured using a perfusion system¹⁰. After freeze-thaw treatment, the synaptosomes were resuspended at 2 mg ml⁻¹ in HBS containing 20 μM pargyline and 0.5 mM ascorbate, and incubated with 0.05 μM $[^3\text{H}]\text{NA}$ at 37 °C for 10 min. Aliquots (1 mg each) of labelled synaptosomes were plated onto glass fibre filters and perfused with 2-ml washes of HBS, each wash delivered at 30-s intervals. After eight washes, the synaptosomes were stimulated with a 30-s wash of HBS containing 30 mM KCl (isotonically substituted for NaCl) in the absence (basal) or presence of 1 mM Ca^{2+} , and this final wash was collected. The radioactivity (d.p.m.) in the collected filtrate and in the filters, solubilized with 1% SDS, was assessed. The fractional efflux (*E*) of $[^3\text{H}]\text{NA}$ was determined for each sample as $E=100 \times \text{filtrate d.p.m.}/(\text{filtrate d.p.m.} + \text{filter d.p.m.})$. The Ca^{2+} -dependent release (*R*) was determined from the difference between *E* found in the presence of Ca^{2+} and *E* found in the absence of Ca^{2+} (basal) and is expressed as means $\pm$ s.e.m. ($n=4$) for successive time intervals.

stimulation by elevated K^+).¹³ In three experiments, the increase in the rate of ionomycin-induced glutamate release averaged $142 \pm 9\%$ (s.e.m.) of control. The introduction of autoactivated kinase into synaptosomes also increased the rate of glutamate release caused by K^+ -induced depolarization in the presence of a physiological concentration (1 mM) of Ca^{2+} (Fig. 1b). In the case of K^+ -induced depolarization, no time lag was observed. In three experiments, the initial rate of K^+ -induced glutamate release was increased to $131 \pm 12\%$ (s.e.m.) of control. The effect of introduced kinase was apparent at physiological (1 mM) concentrations of Ca^{2+} (Fig. 1b), but not low (0.1 mM) concentrations of Ca^{2+} (see Fig. 4).

The K^+ -induced Ca^{2+} -dependent release of [3H]noradrenaline ([3H]NA) was increased to $150 \pm 3\%$ (s.e.m., $n = 4$) of control in synaptosomes preloaded with autoactivated Ca^{2+} /CaM-dependent PK II (Fig. 1c). The increase in the Ca^{2+} -dependent release of [3H]NA was observed within 3 s after K^+ -induced depolarization began (data not shown).

Introduction of Ca^{2+} , CaM and ATP- γ -S without enzyme had no effect on the release of either glutamate or [3H]NA; introduction of enzyme that had not been autoactivated had no consistent effect on neurotransmitter release (data not shown). Also, no significant effect of preloaded autoactivated Ca^{2+} /CaM-dependent PK II on basal (Ca^{2+} -independent) release was observed for either neurotransmitter. The preloaded autoactivated enzyme also had no effect on the uptake or total levels of these neurotransmitters.

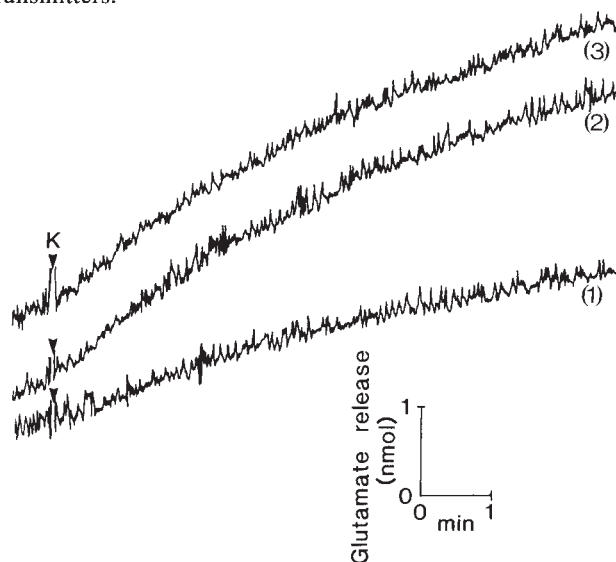


FIG. 2 A Ca^{2+} /CaM-dependent PK II inhibitory peptide (CKIP) decreases release of endogenous L-glutamate from rat brain synaptosomes. Synaptosomes were previously loaded (trace 1) or not loaded (trace 3) with CKIP (3 mM) using freeze-thaw introduction as described in Fig. 1 legend. As an additional control, synaptosomes were freeze-thaw treated and then CKIP was added (trace 2). Release of L-glutamate in the presence of 1 mM Ca^{2+} was measured as described in Fig. 1 legend. The sequence of CKIP is MHRQEAVDCLKKFNARRKLKGA (single-letter amino-acid code), encompassing Met 281 to Ala 302 of the α -subunit of Ca^{2+} /CaM-dependent PK II (see ref. 15), with the autophosphorylation site Thr 286 substituted by alanine. The inhibition constant, K_i , of CKIP for Ca^{2+} /CaM-dependent PK II using synapsin I as substrate was determined using a standard phosphocellulose assay as described elsewhere¹⁶, and was found to be $2.5 \mu M$. CKIP also inhibits the thiophosphorylated autoactivated form of the enzyme (data not shown). This form of the enzyme does not depend on calmodulin for its activity, indicating that CKIP inhibits the enzyme through a mechanism other than that of binding calmodulin. Also, CKIP does not appreciably inhibit Ca^{2+} /CaM-dependent PK III²⁹ (data not shown), further indicating that it is not a general calmodulin antagonist. Although CKIP inhibits Ca^{2+} /CaM-dependent PK I, this kinase phosphorylates only a few substrates and only at sites also phosphorylated by cyclic AMP-dependent protein kinase³⁰. CKIP weakly inhibits protein kinase C. KCl (K; 30 mM) was added when indicated (arrowhead). All traces are to the same scale but are displaced vertically for clarity.

In several experiments, introduction of the catalytic subunit of cyclic AMP-dependent protein kinase into synaptosomes had no effect on Ca^{2+} -dependent [3H]NA release. Consistent with this finding, treatment with 10 μM forskolin, to raise intrasynaptosomal cyclic AMP levels, had no effect on the release of glutamate or [3H]NA (see also ref. 14).

As an alternative way of modulating Ca^{2+} /CaM-dependent PK II activity in synaptosomes, we introduced a selective peptidic inhibitor of the kinase (Ca^{2+} /CaM-dependent PK II inhibitory peptide; CKIP). The sequence of this peptide is based on the regulatory domain of Ca^{2+} /CaM-dependent PK II¹⁵. Introduction of CKIP decreased the rate of glutamate release caused by K^+ -induced depolarization to 50% of control (Fig. 2). In five experiments, the rate of K^+ -induced release of glutamate was decreased to $46 \pm 5\%$ (s.e.m.) of control. Addition of CKIP after freeze-thaw treatment of the synaptosome preparation had

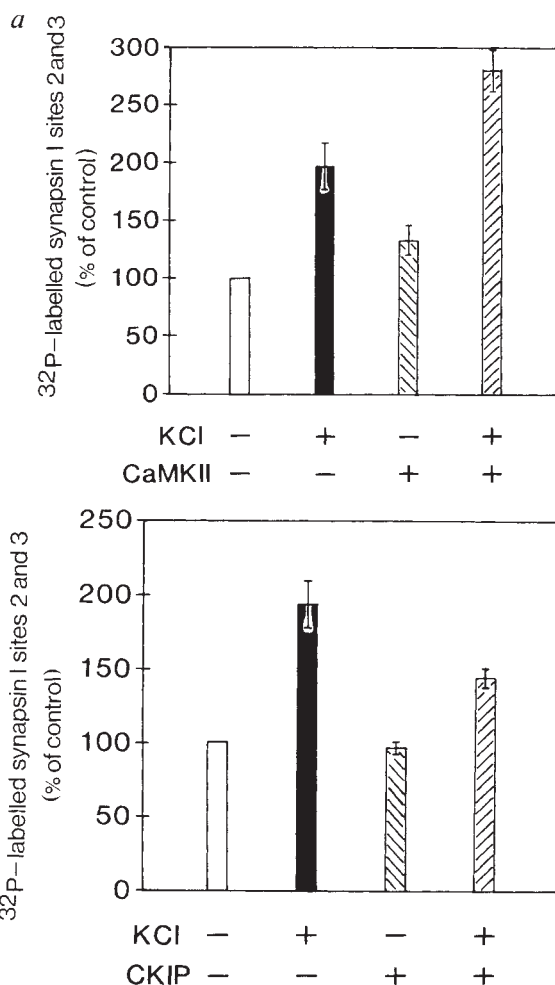
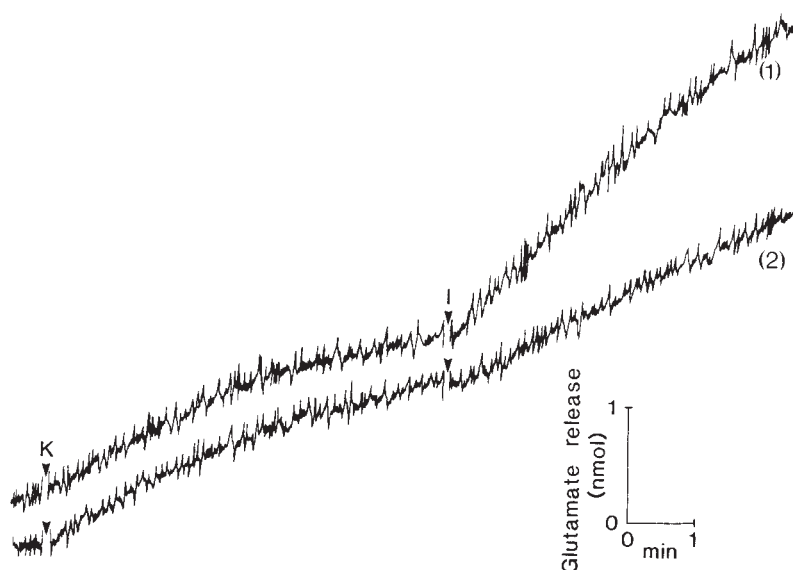


FIG. 3 Effect of freeze-thaw introduction of Ca^{2+} /CaM-dependent PK II (CaMKII; a) or CKIP (b) on ^{32}P -labelling of sites 2 and 3 of endogenous synapsin I in synaptosomes. Synaptosome preparations were resuspended in HBS and loaded or not loaded (1.6 ml final volume), as described in the legend to Fig. 1, with autophosphorylated Ca^{2+} /CaM-dependent PK II (100 μg purified enzyme¹⁶) or with CKIP (3 mM). After washing by centrifugation, the synaptosomes were resuspended in HBS containing 1 mM Ca^{2+} , labelled for 30 min at 37 $^{\circ}C$ with [^{32}P]orthophosphate (2 mCi ml⁻¹, final), and incubated for 15 s without or with 30 mM KCl. The incubations were terminated and the samples analysed using SDS-PAGE³¹. ^{32}P -labelled synapsin I was located in the gels using autoradiography, excised, and subjected to partial proteolytic peptide mapping in 15% SDS-polyacrylamide gels using *Staphylococcus aureus* V8 protease³². The ^{32}P -labelled phosphopeptide of relative molecular mass 35,000 containing sites 2 and 3 of synapsin I^{17,32} was located in the gels and its radioactivity assessed. Values are expressed as percentage of control (unstimulated and not preloaded) and are means $\pm$ s.e.m. ($n = 4$).

FIG. 4 Ca^{2+} /CaM-dependent PK II augments release of endogenous L-glutamate from rat brain synaptosomes after an initial period of release at a low rate. Synaptosomes were previously loaded (trace 1) or not loaded (trace 2) with autoactivated Ca^{2+} /CaM-dependent PK II (200 μg) as described in Fig. 1 legend. Release of L-glutamate, measured as described in Fig. 1 legend, was initiated by addition of elevated K^+ in the presence of 0.1 mM Ca^{2+} . After an initial period of release, ionomycin was added. KCl (K; 30 mM) or 15 μM ionomycin (I) was added when indicated (arrowheads). Both traces are to the same scale but are displaced vertically for clarity.



no effect on subsequent glutamate release (Fig. 2, trace 2). As a separate control, introduction of a peptide with the same amino-acid sequence as CKIP (see Fig. 2 legend) except that it contains threonine at position 286, and is a much less potent inhibitor of Ca^{2+} /CaM-dependent PK II, had only a small effect on K^+ -induced glutamate release (data not shown).

Introduction of autothiophosphorylated Ca^{2+} /CaM-dependent PK II increased the ^{32}P -labelling of synapsin I, a synaptic vesicle-associated phosphoprotein⁹ that is a substrate for Ca^{2+} /CaM-dependent PK II¹⁶. This labelling occurred in the sites in synapsin I (sites 2 and 3) that are specifically phosphorylated by Ca^{2+} /CaM-dependent PK II¹⁷ (Fig. 3a). By contrast, introduction of this autoactivated protein kinase had no effect on the labelling of the protein of relative molecular mass 87,000 that is a specific substrate for protein kinase C¹⁸. With conditions similar to those described for measurement of neurotransmitter release, therefore, introduction of autothiophosphorylated Ca^{2+} /CaM-dependent PK II leads to phosphorylation of a synaptosomal protein that is a specific substrate for Ca^{2+} /CaM-dependent PK II. Although the increase in phosphorylation on introduction of autoactivated kinase was roughly proportional to the increase in neurotransmitter release, we would not necessarily expect these two changes to be linearly related, because neurotransmitter release and its regulation are probably multistep processes. In separate experiments, introduction of CKIP decreased the ^{32}P -labelling of sites 2 and 3 in synapsin I (Fig. 3b).

We have hypothesized^{3,9} that activation of Ca^{2+} /CaM-dependent PK II in response to voltage-dependent or receptor-regulated Ca^{2+} influx into nerve terminals could relieve a constraint on the release of neurotransmitter. The constraint on the release could be due, at least in part, to a tethering of a subpopulation of synaptic vesicles to the cytoskeleton^{9,19}. This subpopulation of vesicles could correspond to a less readily releasable pool of neurotransmitter that has been previously postulated to exist (for example, see ref. 20). In response to the activation of Ca^{2+} /CaM-dependent PK II after Ca^{2+} influx, this pool would be released from the cytoskeleton^{21,22}. These freed synaptic vesicles could then join a different subpopulation of synaptic vesicles, which could correspond to a more readily releasable pool of vesicles available for exocytosis that has been postulated to exist²⁰. On the basis of this two-pool model, we would predict that an initial period of neurotransmitter release, aimed at partially depleting the more readily released pool of vesicles, could magnify the effect of Ca^{2+} /CaM-dependent PK II on neurotransmitter release. Figure 4 shows the results of an experiment designed to test this prediction. We induced an initial period of glutamate release by increasing the concentration of K^+ in the

presence of a low concentration (0.1 mM) of Ca^{2+} . Preloading synaptosomes with autoactivated Ca^{2+} /CaM-dependent PK II had no significant effect on this initial low rate of release (Fig. 4), in contrast to the effect of the kinase on release in the presence of physiological concentrations of Ca^{2+} (Fig. 1b). The subsequent addition of ionomycin to overcome inactivation²³ of the voltage-dependent Ca^{2+} influx increased the rate of glutamate release from synaptosomes preloaded with autoactivated Ca^{2+} /CaM-dependent PK II to 204% of control (Fig. 4). In three experiments, the increase in the rate of release induced by ionomycin after an initial period of K^+ -induced release averaged $195 \pm 7\%$ (s.e.m.) of control.

The study that we have described here demonstrates that Ca^{2+} /CaM-dependent PK II causes an increase in the release of two neurotransmitters—glutamate and noradrenaline—from isolated nerve terminals of mammalian brain. Our work is consistent with the hypothesis that Ca^{2+} /CaM-dependent PK II has a role in mobilizing synaptic vesicles for exocytosis^{3,9}. The results indicate a possible molecular mechanism for the increase in neurotransmitter release resulting from elevated nerve impulse activity²⁴, as well as for the regulation of neurotransmitter release after activation of presynaptic receptors that use Ca^{2+} as a second messenger²⁵. □

Received 17 October; accepted 12 December 1989.

- Nestler, E. J. & Greengard, P. *Nature* **296**, 452–454 (1982).
- Greengard, P., Browning, M. D., McGuinness, T. L. & Llinás, R. *Molecular Mechanisms of Neuronal Responsiveness* (Plenum, New York, 1987).
- Llinás, R., McGuinness, T. L., Leonard, C. S., Sugimori, M. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3035–3039 (1985).
- Llinás, R., Sugimori, M., McGuinness, T., Gruner, J. & Greengard, P. *Biol. Bull.* **169**, 554 (1985).
- Lai, Y., Nairn, A. C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4253–4257 (1986).
- Miller, S. G. & Kennedy, M. B. *Cell* **44**, 861–870 (1986).
- Lou, L. L., Lloyd, S. J. & Schulman, H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9497–9501 (1986).
- Schworer, C. M., Colbran, R. J. & Soderling, T. R. *J. Biol. Chem.* **261**, 8581–8584 (1986).
- De Camilli, P. & Greengard, P. *Biochem. Pharmacol.* **35**, 4349–4357 (1986).
- Nichols, R. A., Wu, W. C.-S., Haycock, J. W. & Greengard, P. *J. Neurochem.* **52**, 521–529 (1989).
- Li, H.-C., Simonelli, P. F. & Huan, L.-J. *Meth. Enzym.* **159**, 346–356 (1988).
- Bass, M., Pant, H. C., Gainer, H. & Soderling, T. R. *J. Neurochem.* **49**, 1116–1123 (1987).
- Sanchez-Prieto, J., Sihra, T. S. & Nicholls, D. G. *J. Neurochem.* **49**, 58–64 (1987).
- Nichols, R. A., Haycock, J. W., Wang, J. K. T. & Greengard, P. *J. Neurochem.* **48**, 615–621 (1987).
- Payne, M. E. *et al.* *J. Biol. Chem.* **263**, 7190–7195 (1988).
- McGuinness, T. L., Lai, Y. & Greengard, P. *J. Biol. Chem.* **260**, 1696–1704 (1985).
- Czernik, A. J., Pang, D. T. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7518–7522 (1987).
- Wu, W. C.-S., Walaas, S. I., Nairn, A. C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5249–5253 (1982).
- Hirokawa, N., Sobue, K., Kanda, K., Harada, A. & Yoritani, H. *J. Cell Biol.* **108**, 111–126 (1989).
- Zucker, R. S. in *Neuromodulation* (eds Kaczmarek, L. K. & Levitan, I. B.) 243–263 (Oxford, New York, 1987).
- Bähler, M. & Greengard, P. *Nature* **326**, 704–707 (1987).
- Sihra, T. S., Wang, J. K. T., Gorelick, F. S. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8108–8112 (1989).
- Nachsen, D. A. & Blaustein, M. P. *J. gen. Physiol.* **76**, 709–728 (1980).
- Magleby, K. L. & Zengel, J. E. *J. gen. Physiol.* **80**, 613–638 (1982).
- Chesselet, M.-F. *Neuroscience* **12**, 347–375 (1984).
- Dunkley, P. R., Jarvie, P. E., Heath, J. W., Kidd, G. J. & Rostas, J. A. P. *Brain Res.* **372**, 115–129 (1986).

27. Nicholls, D. G. & Sihra, T. S. *Nature* **321**, 772-773 (1986).
 28. Nicholls, D. G., Sihra, T. S. & Sanchez-Prieto, J. J. *Neurochem.* **49**, 50-57 (1987).
 29. Nairn, A. C., Baghat, B. & Palfrey, H. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7939-7943 (1985).
 30. Nairn, A. C. & Greengard, P. *J. Biol. Chem.* **262**, 7273-7281 (1987).
 31. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
 32. Huttner, W. B. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5402-5406 (1979).

ACKNOWLEDGEMENTS. We thank Ms Catherine Carlson and Ms Sara Engl for excellent technical assistance, Drs Fred Gorelick and Daniele Piomelli for valuable suggestions and discussions, and Dr Yvonne Lai for the gift of Ca^{2+} -CaM-dependent PK II used in the phosphorylation experiments. Some of this work was performed while R.A.N. held a fellowship with the Norman & Rosita Winston Foundation. This work was supported by the USPHS (P.G.).

Requirement for integration of signals from two distinct phosphorylation pathways for activation of MAP kinase

Neil G. Anderson*, James L. Maller†, Nicholas K. Tonks‡ & Thomas W. Sturgill*§

* Department of Internal Medicine and Pharmacology, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA

† Department of Pharmacology, University of Colorado School of Medicine, Denver, Colorado 80262, USA

‡ Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

MAP kinase (relative molecular mass, 42,000), a low abundance serine-threonine protein kinase, is transiently activated in many cell types by a variety of mitogens, including insulin, epidermal growth factor, and phorbol esters^{1,2}. *In vitro*, MAP kinase will phosphorylate and reactivate S6 kinase II previously inactivated by phosphatase treatment³. Because many of the stimuli that activate MAP kinase are also stimulators of cell proliferation, and regulation of the cell cycle seems to involve a network of protein kinases, MAP kinase could be important in the transmission of stimuli eventually leading to the progression from G0 to G1 in the cell cycle. Activated MAP kinase contains both phosphotyrosine and phosphothreonine⁴. We report here that MAP kinase can be deactivated completely by treatment with either phosphatase 2A, a protein phosphatase specific for phosphoserine and phosphothreonine, or CD45, a phosphotyrosine-specific protein phosphatase. We demonstrate that MAP kinase is only active when both tyrosyl and threonyl residues are phosphorylated and suggest therefore that the enzyme functions *in vivo* to integrate signals from two distinct transduction pathways.

Although phosphatase 2A is widely recognized as a protein phosphatase specific for phosphoserine and phosphothreonine⁵, it has significant phosphotyrosine-specific phosphatase activity towards some substrates⁷. It was not certain, therefore, whether the complete deactivation of MAP kinase that occurs when the enzyme is incubated with phosphatase 2A is caused by dephosphorylation of phosphothreonyl residues or phosphotyrosyl residues, or both. Also, incubation of MAP kinase with another protein phosphoserine-phosphothreonine-specific phosphatase, phosphatase 1, which has no reported protein tyrosine phosphatase activity, does not significantly affect MAP kinase activity³. For these reasons, we performed experiments to correlate changes in MAP kinase activity with changes in phosphoamino-acid content. The availability of a highly purified preparation of a phosphotyrosine-specific protein phosphatase, CD45, enabled us to test directly the role of tyrosine phosphorylation in controlling MAP kinase activity.

CD45 is a member of a family of glycoproteins that span the membranes of most haematopoietic cells⁸. The cytoplasmic domain of CD45 possesses protein phosphatase activity specific

for phosphotyrosyl residues⁸. Consequently, CD45 is hypothesized to participate in a mechanism of signal transduction in these cells involving protein tyrosine dephosphorylation. CD45 has no sequence similarity with serine-threonine phosphatases⁹, and does not cause the release of phosphate from any of several serine-threonine-phosphorylated protein substrates (data not shown).

Figure 1 shows that MAP kinase, purified from fibroblasts treated with epidermal growth factor, was completely inactivated by incubation with either phosphatase 2A (10 U ml⁻¹) or CD45 (10 U ml⁻¹). Loss of activity was rapid, and was essentially complete after 20-40 min in either case. Inactivation of MAP kinase was attributable to the phosphatases used and not to contaminating proteases, for the following reasons. First, treatment was performed in the presence of protease inhibitors. Second, no shift in the mobility of the [³²P]phosphoprotein of relative molecular mass 42,000 (*M*_r42K) was detectable after phosphatase treatment (data not shown). Finally, control incubations of MAP kinase with either phosphatase inactivated by prior incubation with specific inhibitors (fluoride-EDTA for phosphatase 2A, and vanadate for CD45) resulted in minimal loss of MAP kinase activity even when incubated for much longer times, up to 1 h (Fig. 1). The latter observation also indicates that there was no thermal denaturation.

The preparation of CD45 used was purified by antibody-affinity chromatography⁸ and lacked serine-threonine phosphatase activity. Inclusion of okadaic acid (1 µM) has no effect on the inactivation of MAP kinase by CD45 (data not shown).

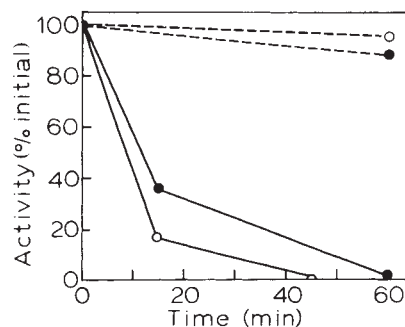
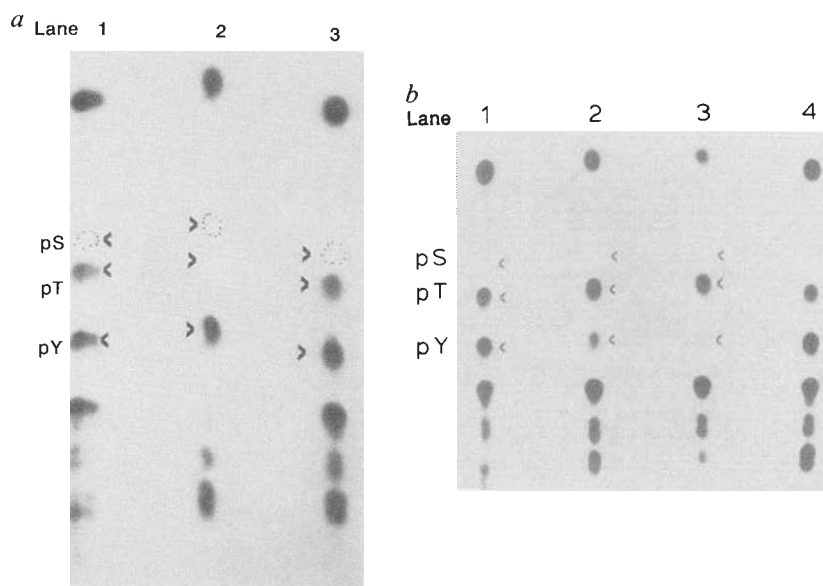


FIG. 1 Deactivation of MAP kinase by phosphatase 2A and CD45. MAP kinase activity was measured after treatment at 30 °C with phosphatase 2A (10 U ml⁻¹, ○—○) or CD45 (10 U ml⁻¹, ●—●). Control incubations (○—○, ●—●) were performed with inactivated phosphatases.

METHODS. MAP kinase was purified from 3T3-L1 fibroblasts treated with epidermal growth factor (30 nM, 10 min). MAP kinase, in Superose-12 buffer additionally containing BSA (1 mg ml⁻¹) and protease inhibitors³, was treated with either of the phosphatases at 30 °C in a final volume of 50 µl. At each time point, aliquots were removed and the phosphatase reaction stopped by incubation with inhibitors (10 mM NaF-1 mM EDTA for phosphatase 2A; 200 µM sodium vanadate for CD45; all final concentrations) for 10 min. MAP kinase activity was then measured using myelin basic protein (Sigma) as previously described¹⁰ with 5 c.p.m. fmol⁻¹ [³²P]ATP; myelin basic protein is an alternative substrate for MAP kinase (data not shown). Initial activity incorporated into myelin basic protein was 138,000 and 35,000 c.p.m. per 10 min for the CD45 and 2A experiments shown, respectively. Control incubations were performed using phosphatases inactivated by prior incubation with inhibitors. Data shown for inactivation by phosphatase 2A and CD45 are from two representative experiments combined into one graph; the points are averages of duplicates that differed by <5%. The preparation of the catalytic subunit of phosphatase 2A was as previously reported³; CD45 had been purified from human spleen to apparent homogeneity⁸. One unit of phosphatase catalyzes the release of 1 nmol min⁻¹ of phosphate from [³²P]phosphorylase a or [³²P]phosphotyrosyl-lysozyme⁹, for phosphatase 2A and CD45, respectively. Phosphatase 2A was in 50 mM Tris buffer, pH 7.0, 0.1 mM EGTA, 50 mM β-mercaptoethanol and 60% glycerol; CD45 was in 25 mM imidazole, pH 7.2, 15 mM β-mercaptoethanol, 0.1 mM EGTA, 0.05% Triton X-100 and 20% glycerol. Dilutions of phosphatase into reaction mixtures were 50- and 20-fold for phosphatase 2A and CD45, respectively.

§ To whom correspondence should be addressed.

FIG. 2 Phosphoamino-acid analysis of [32 P]MAP kinase before and after deactivation with phosphatases. **a**, [32 P]MAP kinase treated with phosphatase 2A for 0 min (lane 1) or 15 min (lane 2) or with inactivated phosphatase for 60 min (lane 3). **b**, [32 P]MAP kinase treated with CD45 for 0, 15 and 60 min (lanes 1–3, respectively) or with inactivated phosphatase for 60 min (lane 4). Phosphoserine, pS; phosphothreonine, pT; phosphotyrosine, pY. **METHODS.** [32 P]MAP kinase was purified from 3T3-L1 cells labelled¹ for 2 h at 37 °C with 32 Pi (1 mCi ml⁻¹) followed by treatment with epidermal growth factor (30 nM, 10 min). The enzyme was incubated with phosphatases exactly as described in Fig. 1. Aliquots were removed and assayed for MAP kinase activity after stopping the activities of the phosphatases with inhibitors. The activities of the remainder of the reaction mixtures were stopped by addition of SDS sample buffer and boiled for 3 min. Phosphoamino-acid analysis after partial acid hydrolysis was performed on the 32 P-labelled bands after transfer to Immobilon P (Millipore) using procedures described by Kamps and Sefton¹¹. Phosphoamino acids were separated on cellulose TLC plates (Merck) by electrophoresis (60 min, 1 kV) in buffer of pH 3.5.



Okadaic acid, at this concentration, completely inhibits phosphatases 1, 2A and 2B (ref. 6). The specificity of CD45 for phosphotyrosine, with these additional controls, strongly indicates that the phosphotyrosyl site or sites in MAP kinase are regulatory.

Further experiments showed that the changes in activity of MAP kinase caused by the phosphatases were correlated with changes in phosphoamino-acid content. The results of phosphoamino-acid analyses are shown in Fig. 2. [32 P]MAP kinase contained exclusively phosphotyrosine and phosphothreonine, as previously demonstrated (lane 1, Fig. 2a; lane 1, Fig. 2b). Incubation of [32 P]MAP kinase with either inactivated phosphatase 2A or inactivated CD45, which does not alter activity, did not affect this composition (Fig. 2a, lane 3; Fig. 2b, lane 4). Treatment with active phosphatase 2A caused complete deactivation of MAP kinase, like that described above, that correlated with the complete removal of phosphothreonine without appreciable alteration of the phosphotyrosyl content of MAP kinase (Fig. 2a, lane 2). Treatment with the protein tyrosine phosphatase CD45 for 60 min caused complete loss of kinase activity, correlated with essentially complete removal of phosphotyrosine without appreciable change in phosphothreonine content (Fig. 2b, lane 3). Partial dephosphorylation of phosphotyrosine by treatment with CD45 for 15 min resulted in a proportional loss of activity (Fig. 1, Fig. 2b, lane 2). These findings were verified in repeated experiments.

Incubation with either phosphatase 2a or CD45 twice as long as required to deactivate MAP kinase does not result in further removal of 32 PO₄, and simultaneous incubation with both phosphatases is required to release all of the 32 PO₄ from [32 P]MAP kinase (data not shown). These observations support the conclusion that each phosphatase is catalytically active towards different sites. Because of the differential recoveries of each phosphoamino acid after acid hydrolysis, we cannot yet estimate the likely proportions of phosphothreonine and phosphotyrosine in MAP kinase.

The simplest interpretation of the data obtained is that both threonine and tyrosine phosphorylations must occur for activation of MAP kinase. Neither of these phosphorylations alone seem to be sufficient for activation. This requirement for the phosphorylation of two different species of acceptor amino acid for activation of a kinase is novel. The converse situation exists for activation of maturation promotion factor (MPF)¹². The 34K kinase subunit of MPF, the *cdc2* gene product, is also phos-

phorylated during interphase on both tyrosyl and threonyl residues¹². Dephosphorylation of the phosphotyrosyl site(s) in the *cdc2* gene product does not result in activation of MPF activity¹³, indicating that dephosphorylation of both species of phosphoamino acid is required for activation. The *cdc2* gene product is required in yeast (and presumably in eukaryotic cells) for both the G1-S and the G2-M phase transitions¹². Regulation of enzyme activity by phosphorylation of both tyrosine and threonine residues could reflect a more general control mechanism than previously appreciated.

The protein tyrosine phosphatase of low *M_r* from placenta (termed 1B; 23 U ml⁻¹ final concentration), which has sequence similarity with CD45, is dramatically less effective than CD45 (at 10 U ml⁻¹) in deactivating and dephosphorylating MAP kinase (data not shown). By contrast, protein tyrosine phosphatase 1B is 20- to 50-fold more active than CD45 towards [32 P]phosphotyrosyl-lysozyme⁸. Differences in substrate specificity for phosphotyrosyl phosphatases are expected to be important for cellular regulation.

A requirement for phosphorylation of both tyrosyl and threonyl sites for activity indicates that MAP kinase could serve to integrate information from converging signal transduction pathways. This view is supported by evidence equating active MAP kinase with one form of pp42 (specifically, pp42A)², widely studied as a principal target for tyrosine phosphorylation in quiescent cells stimulated with several mitogens, including phorbol esters (see ref. 2). We hypothesize that the downstream target or targets for MAP kinase, potentially including S6 kinase^{3,14}, regulate one or more steps in the cell cycle that are common to the actions of the several mitogens that stimulate MAP kinase activity. The report here that a phosphotyrosine-specific phosphatase dephosphorylates and deactivates MAP kinase *in vitro* will now facilitate the finding of at least two MAP-kinase kinases able to phosphorylate and reactivate the deactivated enzyme. Elucidating the pathway of activation of MAP kinase should provide insights into the mechanism of mitogenic signalling and cell-cycle control. □

Received 29 September; accepted 30 November 1989.

1. Ray, L. B. & Sturgill, T. W. *J. biol. Chem.* **263**, 12721–12727 (1988).
2. Rossomando, A. J., Payne, D. M., Weber, M. J. & Sturgill, T. W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6940–6943 (1989).
3. Sturgill, T. W., Ray, L. B., Erikson, E. & Maller, J. L. *Nature* **334**, 715–718 (1988).
4. Ray, L. B. & Sturgill, T. W. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3753–3757 (1988).

5. Ray, L. B. & Sturgill, T. W. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1502-1506 (1987).
6. Cohen, P. A. *Rev. Biochem.* **58**, 453-508 (1989).
7. Jones, S. W., Erikson, R. L., Ingebritsen, V. M. & Ingebritsen, T. S. *J. biol. Chem.* **264**, 7747-7753 (1989).
8. Tonks, N. K., Charbonneau, H., Diltz, C. D., Fischer, E. H. & Walsh, K. A. *Biochemistry* **27**, 8695-8710 (1988).
9. Charbonneau, H., Tonks, N. K., Walsh, K. A. & Fischer, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7182-7186 (1988).
10. Cicirelli, M. F., Pelech, S. L. & Krebs, E. G. *J. biol. Chem.* **263**, 2009-2019 (1988).
11. Kamps, M. P. & Sefton, B. M. *Analyt. Biochem.* **176**, 22-27 (1989).
12. Murray, A. W. & Kirschner, M. W. *Science* **246**, 614-621 (1989).
13. Morla, A. O., Draetta, G., Beach, D. & Wang, J. Y. J. *Cell* **58**, 193-203 (1989).
14. Gregory, J. S., Boulton, T. G., Sang, B. & Cobb, M. H. *J. biol. Chem.* **264**, 18397-18401 (1989).

ACKNOWLEDGEMENTS. We thank L. Bryan Ray for performing preliminary phosphoamino-acid analyses of MAP kinase inactivated by phosphatase 2A. This work was supported by the US Public Health Service, the American Cancer Society, and the Juvenile Diabetes Foundation, Int. We thank Michael Weber for reading the manuscript and for discussions.

Low dimensional chaos in cardiac tissue

Dante R. Chialvo, Robert F. Gilmour Jr* & Jose Jalife

Department of Pharmacology, SUNY Health Science Center,
766 Irving Avenue, Syracuse, New York 13210, USA

* Department of Physiology, New York State College of Veterinary
Medicine, Cornell University, Ithaca, New York 14853-6401, USA

CHAOS is a term used to characterize aperiodic activity arising in a dynamical system, or in a set of equations describing the system's temporal evolution as a result of a deterministic mechanism that has sensitive dependence on initial conditions¹. Chaos, in that sense, has been proposed to make an important contribution to normal and abnormal cardiac rhythms²⁻⁶. To date, however, descriptions of chaos in heart tissue have been limited primarily to periodically forced cardiac pacemakers^{2,6}. Because many cardiac rhythm disturbances, particularly those initiated or perpetuated by re-entrant excitation, originate from within non-pacemaker cardiac tissues⁷⁻⁹, demonstrations of chaos in non-pacemaker tissue might provide a deterministic explanation for a wide variety of complex dysrhythmias. Here we report experimental evidence for chaotic patterns of activation and action potential characteristics in externally driven, non-spontaneously active Purkinje fibres and ventricular muscle. The results indicate that there is an apparent link between the mechanism of low dimensional chaos and the occurrence of reflected responses which could lead to more

spatially disorganized phenomena. A detailed mechanism for the low dimensional chaos observed experimentally is pursued using a difference equation model. Critical features of the model include a non-monotonic relationship between recovery time during rhythmic stimulation and the state of membrane properties, and a steeply sloped recovery of membrane properties over certain ranges of recovery times. Besides explaining our results, the analytical model may pertain to irregular dynamics in other excitable systems, particularly the intact dysrhythmic heart.

Chaos has been demonstrated in several experimental and theoretical models of periodically forced oscillators, including cardiac pacemakers^{1,2,6}. Although non-autoactive excitable media may also display chaotic dynamics¹⁰, aperiodic responses of forced non-pacemaker cardiac tissue have been observed only under restricted experimental conditions^{11,12}, which include low amplitude of stimulation in conjunction with supernormality.

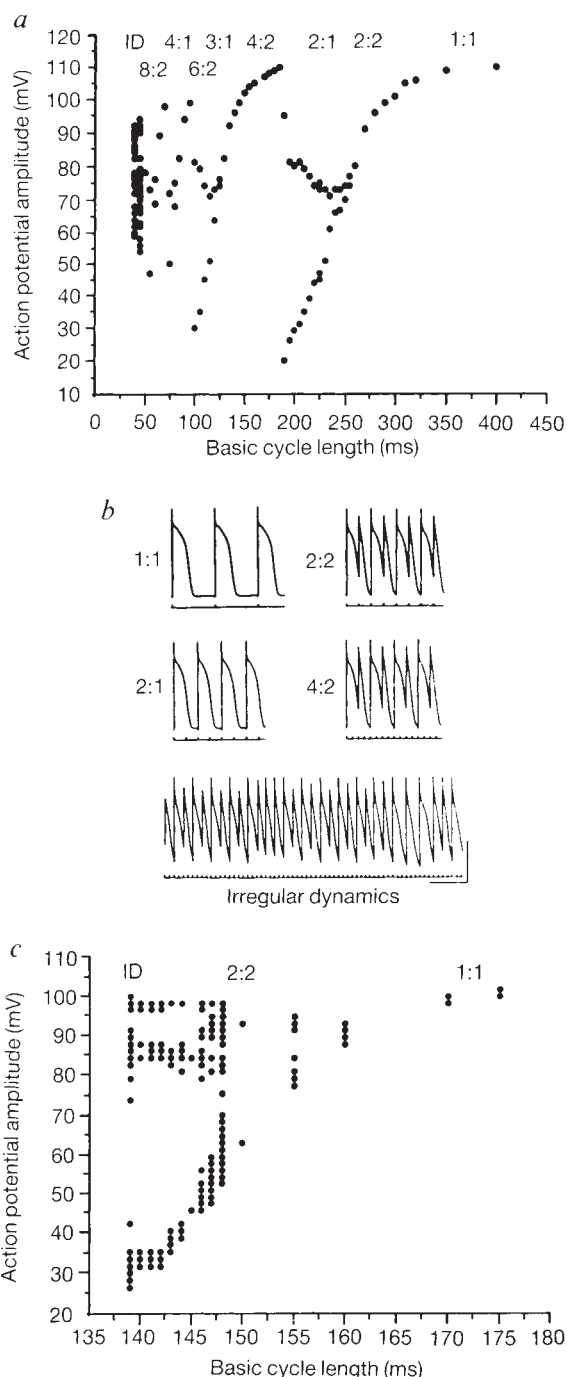


FIG. 1 Period-doubling bifurcations and chaotic patterns of action potential amplitude in cardiac Purkinje fibres. *a*, Bifurcation diagram showing the relationship between pacing basic cycle length and beat-to-beat action potential amplitudes in a sheep cardiac Purkinje fibre. Stimulus:response locking indicated at top (ID, irregular dynamics). Stimulus intensity, $1.5 \times$ threshold; KCl concentration, 2.7 mM. *b*, Analog recording of bifurcations and irregular dynamics obtained in a dog cardiac Purkinje fibre. Upper trace is transmembrane action potential and lower trace is stimulus record 1:1, 2:2, 2:1 and 4:2 locking were recorded at basic cycle lengths (BCLs) of 500, 145, 140 and 75 ms, respectively. Irregular dynamics preceding 3:1 locking were detected at BCL of 53 ms. Stimulus intensity, $2.0 \times$ threshold; KCl concentration, 2.7 mM; calibration bars are 50 mV and 500 ms. *c*, Cascade of period doubling bifurcations, starting at 1:1 locking, leading to irregular dynamics in a sheep cardiac Purkinje fibre. Stimulus intensity, $2.6 \times$ threshold; KCl concentration, 4.0 mM.

METHODS. Free running false tendons were obtained from the left ventricles of adult sheep ($n=13$) following electrocution or from adult dogs ($n=14$) following anaesthesia using pentobarbital sodium (35 mg kg^{-1} intravenously). The preparations were mounted in a tissue bath and superfused with normal Tyrode solution maintained at pH 7.4, $\text{PO}_2 > 400 \text{ mm Hg}$ and $T=37.0^\circ\text{C}$. The composition of the Tyrode solution (in mM) was: MgCl_2 0.5, NaH_2PO_4 0.9, CaCl_2 2.0, NaCl 137.0, NaHCO_3 24.0, KCl 2.7 or 4.0 and glucose 5.5. A concentration of 2.7 mM was used for KCl in some experiments to enhance supernormality^{11,22}. Transmembrane recordings were obtained using floating glass microelectrodes, as described previously^{12,27}. Pacing stimuli were 2.0 ms duration rectangular pulses obtained from a pulse generator and delivered through an isolation transformer at variable intensities.

Further studies of a simple experimentally based mathematical model¹² indicate that chaos might be a more general phenomenon associated with particular modes of nonlinear, time-dependent recovery of critical membrane properties in excitable tissues. Specifically, chaotic dynamics are predicted under experimental circumstances in which either significant delay to activation at very early recovery intervals, or non-monotonic recovery of action potential duration occurs. These theoretical predictions are explored here by a systematic examination of aperiodic behaviour in non-spontaneously active cardiac Purkinje fibres and ventricular muscle.

To demonstrate low dimensional chaotic behaviour in cardiac tissue, we initially determined whether a cascade of period-doubling bifurcations leading to more irregular activity occurred during decremental pacing. Figure 1a illustrates initial period doubling bifurcations of action potential amplitude occurring in a cardiac Purkinje fibre paced at increasing frequencies (that is, at decreasing cycle lengths). Period-doubling bifurcations were associated with changes in stimulus:response locking, in that 1:1 locking was replaced by 2:2, 2:1 by 4:2, 3:1 by 6:2, and 4:1 by 8:2, until irregular activity arose at very brief cycle lengths. Careful exploration of the transition between 4:2 and 3:1 revealed that irregular dynamics also occurred at intermediate pacing rates. An example of such a transition is shown in Fig. 1b. The results shown in Fig. 1a and b provide experimental evidence for the theoretical predictions made previously¹².

In Fig. 1c, stimulus amplitude is higher than that used in Fig. 1a and b. As the basic cycle length was abbreviated, a 2:2 locking pattern was followed by an additional period-doubling bifurcation (that is, 4:4) and then by irregular dynamics. The structure of the bifurcation diagram is similar to that observed during the period doubling route to chaos in unidimensional maps¹³⁻¹⁶. We sought additional experimental evidence to demonstrate conclusively that this irregular activity was in fact deterministic chaos. A return map of irregular sequences of action potential amplitude is depicted in Fig. 2. Successive data points are confined to specific regions of the return map, rather

than being distributed randomly, indicating an underlying one-dimensional deterministic mechanism. Moreover, the data are distributed along two segments: a long descending segment that intersects the identity line (that is, $APA_{n+1} = APA_n$, where APA is the action potential amplitude) and a shorter ascending segment. Both the slope of the descending segment (less than minus one) and the order of appearance of consecutive data in the map (Fig. 2b) indicate a sensitive dependence on initial conditions and, therefore, a deterministic (chaotic) origin of the irregular activity¹²⁻²⁰.

A simple theoretical model originally introduced by Guevara *et al.*¹⁸ and further modified by us¹² can account for these experimental observations. As summarized in Fig. 3a, the main arguments used in the model are: (1) the recovery of relevant membrane properties after each action potential is nonlinear and (2) the fixed time between pulses during rhythmic stimulation generates reciprocal interactions between the recovery time (represented in the model by diastolic interval) and the degree of recovery of membrane properties (latency and action potential duration). In such a feedback loop, there is a particular range of pacing cycle lengths within which period-doubling bifurcation and chaos are expected.

Figure 3b shows the predicted bifurcation sequence for experiments similar to those illustrated in Figs 1c and 2. In both experimental and theoretical cases, chaos always occurred after at least one period-doubling bifurcation. Likewise, the experimental and theoretically derived return maps during chaos (Figs 2 and 3c, respectively) were similar. It is apparent from the analytical model that the slope of the descending segment of the map is determined by the time constant of the fast component of the time to full repolarization function^{21,22} (see the time to full repolarization function in Fig. 3a for a diastolic interval between -100 and 0 ms), whereas the ascending branch reflects prolongation of the time to full repolarization at short (that is, negative) diastolic intervals (Fig. 3a). Supernormal excitability²³ facilitates activation at short diastolic intervals, where recovery of action potential amplitude and time to full repolarization are

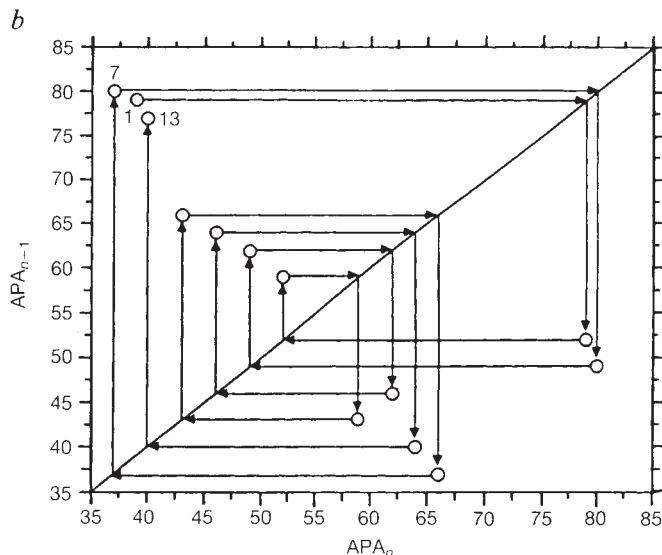
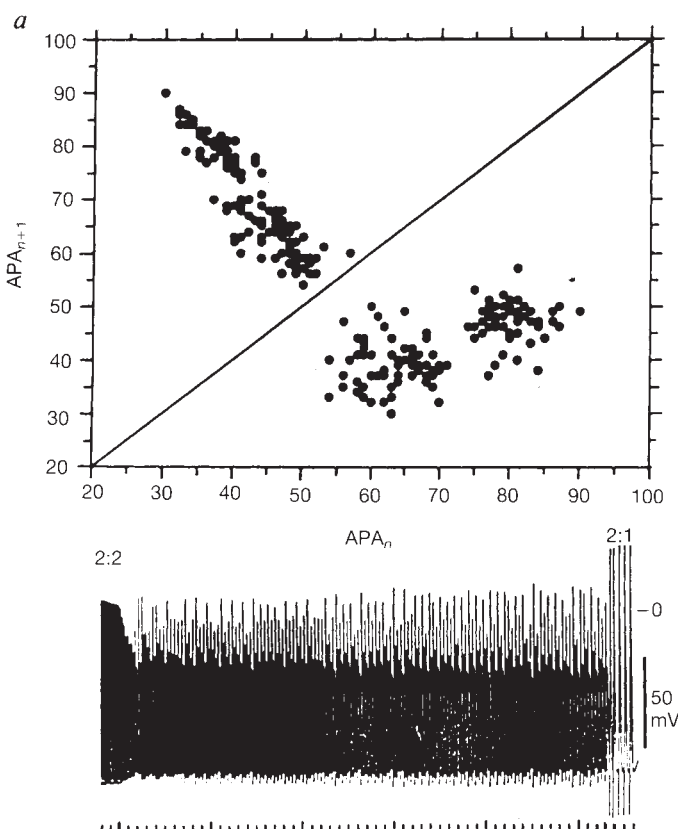


FIG. 2 Low dimensional chaos in a sheep cardiac Purkinje fibre. a, Lower record, Transmembrane recording obtained during pacing at a BCL of 180 ms. Brief records during pacing at BCL of 185 ms (2:2 locking, left) and 175 ms (2:1 locking, right) are also shown. Stimulus intensity, $2.0 \times$ threshold; KCl concentration, 4.0 mM. Timescale is 1 s per division. Upper panel, Corresponding one-dimensional return map for APA, where APA_{n+1} is given for each preceding APA_n . For $APA_n < 75$ mV, APA_n and APA_{n+1} are linearly related (slope, 1.441; $r^2 = 0.891$). b, Return map of a portion of the irregular dynamics shown in a. Arrows indicate the sequence of the map, starting at 1 and ending at 13 (upper left). The fibre is different from that used in Fig. 1c. Methods are described in the legend to Fig. 1.

non-monotonic ('notched') (Fig. 3a). As the rate of stimulation is increased, the slope of the map (Fig. 3c) at the intersection with the identity line becomes progressively steeper until it equals one, at which point a period-doubling bifurcation occurs. Additional increases in stimulus rate produce a cascade of period-doubling bifurcations and chaos. These observations reflect the known dynamical consequences for functions that contain both a steep slope and a minimum value^{13,24,25}, so the existence of a minimum in the map guarantees that higher period doubling and chaos will occur.

In the absence of a minimum, represented here by the inflection point in the time to full repolarization versus diastolic interval function, 1:1 locking is replaced only by an initial period doubling bifurcation (2:2) (refs 18, 24) and then by 2:1 locking. This particular bifurcation scenario (presented in Fig. 1a and b) can be readily reproduced in the model by lowering stimulation amplitude so that the time to full repolarization function minimum disappears¹².

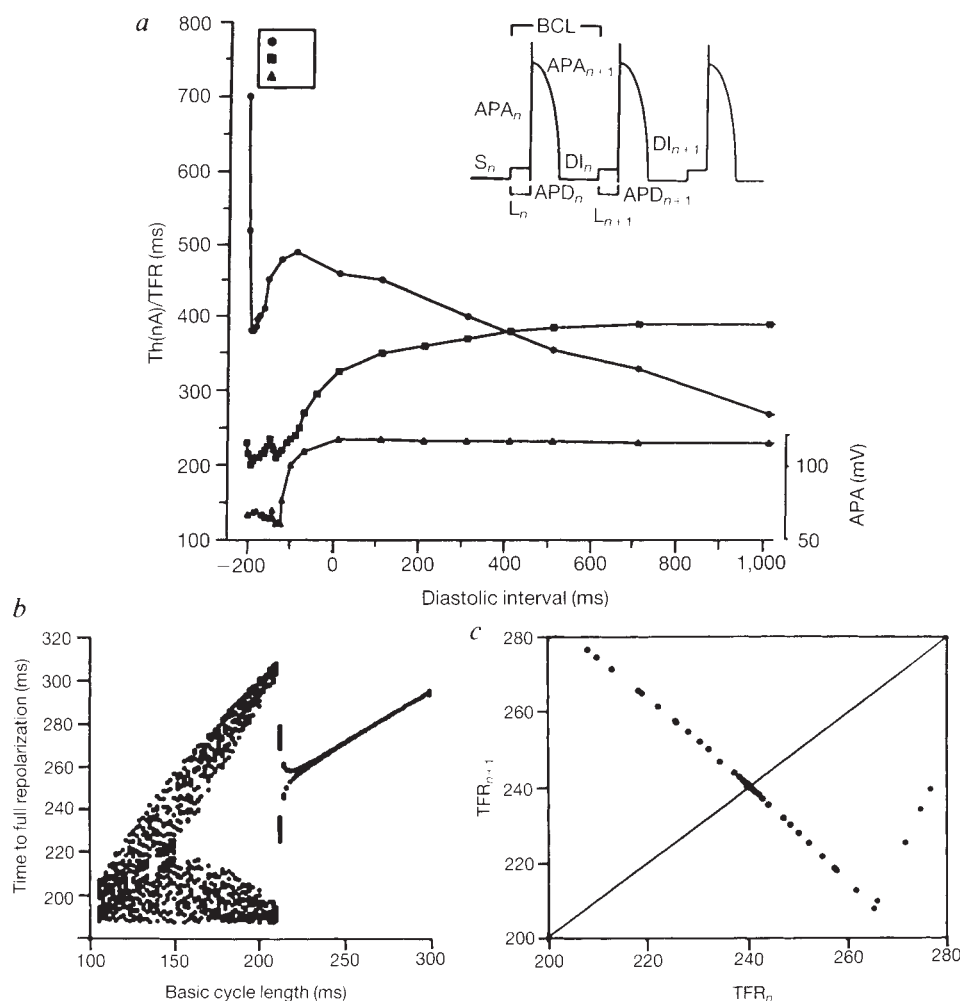
Period-doubling bifurcations of action potential characteristics and chaotic patterns of activation were also demonstrated in ventricular muscle cells of Purkinje-muscle preparations (Fig. 4). Rhythmic stimulation of the Purkinje fibre resulted in bifurcation behaviour in ventricular muscle cells (Fig. 4a) very similar to that observed in the Purkinje fibre preparation. Return maps of ventricular muscle responses reconstructed during periodic (period three in Fig. 4b) and aperiodic (Fig. 4c) activity indicate

the existence of the same dynamical mechanism (steep slope and minimum) already described for Purkinje fibres. In ventricular muscle, as in the Purkinje fibres, the critical slope in the return map (Fig. 4c) is provided by the fast component of the action potential duration restitution function. The ascending segment of the map, reflecting significant conduction delays through Purkinje-muscle junctional cells following activation at short diastolic intervals, is associated with depressed conduction (induced by heptanol and hyperkalaemia^{26,27}). Under these experimental conditions, the 'notches' in the time to full repolarization curve (Fig. 3a) are accentuated, thereby permitting the development of complex dynamics.

Whereas the chaotic dynamics in Purkinje fibres occurs in relatively homogeneous tissue (no spatial component), an important spatial parameter is present in the Purkinje-muscle preparations. A fascinating point is that both situations (essentially space-independent in Purkinje fibres, but space-dependent in Purkinje-muscle preparations) collapsed in a one-dimensional behaviour that is reproduced by a simple model.

For the reasons given, induction of chaotic activation at Purkinje-muscle junctions depends significantly on activation delay. As expected, prolonged activation delay also facilitates reflection^{27,28} of impulses between ventricular muscle and the Purkinje fibre and within ventricular muscle (Fig. 4d). Interestingly, both reflection and chaos occur over the same range of cycle lengths, whereas period two is commonly observed in the

FIG. 3 Period-doubling bifurcations and chaos in an analytical model of cardiac cells. **a**, Inset (upper right) shows the relevant parameters for the difference equation model. If latency to activation (L), total action potential duration (APD), current threshold for all-or-none excitation (Th) and action potential amplitude (APA) are functions (F , G , Z and Q , respectively) of the diastolic interval (DI), where DI_n is $BCL - L_n - APD_n$, S is stimulus and APA is the difference between the resting membrane potential and the peak voltage attained following a stimulus, then (1) $L_{n+1} = F(DI_n)$; (2) $APD_{n+1} = G(DI_n)$; (3) $Th_{n+1} = Z(DI_n)$; and (4) $APA_{n+1} = Q(DI_n)$. No distinction has been made between active and passive membrane responses. Recovery of Th (circles), time to full repolarization (TFR, where TFR is $L + APD$; squares), and APA (triangles) are given as functions of DI . Values were obtained experimentally from a sheep Purkinje fibre, as described previously²⁷. For the simulation presented in **b** and **c**, model functions F , G and Z were reconstructed by linear interpolation between data points. As measurements of L and APD during stimulation at high frequencies involve imprecise estimations of the completion of repolarization (see, for example, the analog records of irregular dynamics in Fig. 1b), APA was used to characterize the experimental dynamics. APA has a defined functional relationship with DI and, therefore, with TFR (see **a**). For presentation of the numerical results, TFR was chosen because it is the relevant parameter in the model. Note the supernormal phase of excitability (when Th is less than expected from a monotonic exponential decay) at DI of -180 to -20 ms. Note also non-monotonic behaviour of TFR and APA (notches) at DI between -200 and -180 ms and between -145 and



-105 ms. **b**, Predicted bifurcation diagram showing the relationship between BCL and TFR for 60 iterations of the model at each BCL. **c**, Return map for TFR at a BCL of 150 ms.

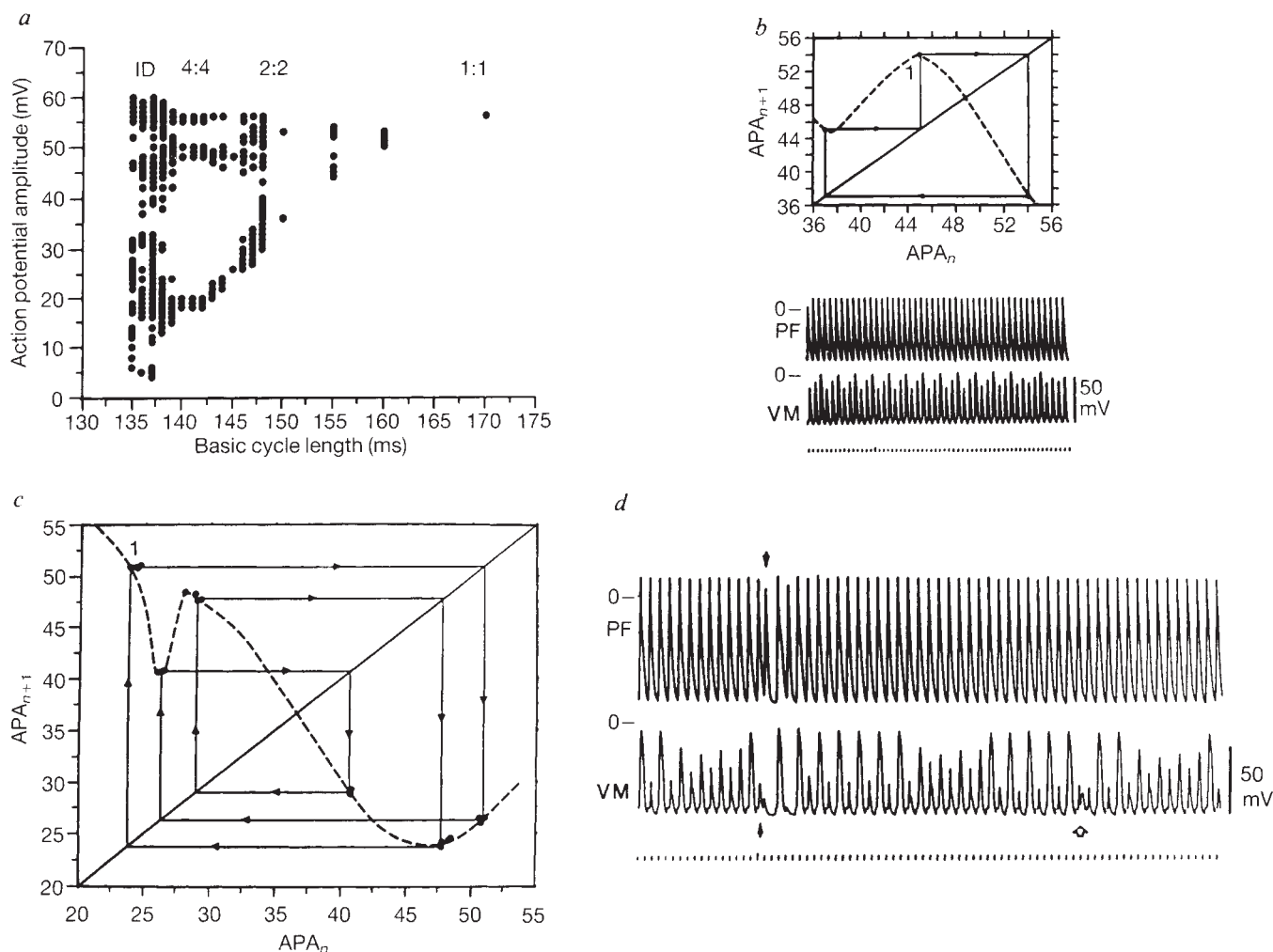


FIG. 4 Period-doubling bifurcations and chaos in canine Purkinje-muscle preparations. *a*, Bifurcation diagram showing beat-to-beat ventricular muscle action potential amplitude at different pacing cycle lengths. Similar results were obtained in five other preparations. *b*, Lower record, Transmembrane recordings of period 3 in ventricular muscle (VM; lower trace) during pacing of Purkinje fibre (PF; upper trace) as a BCL of 135 ms. Timescale 100 ms per division. Upper record, Corresponding return map for VM APA. *c*, Return map for aperiodic VM APA at a BCL of 147 ms. Four consecutive sequences are shown, starting at 1 (upper left). When all data were considered ($n=62$), the slope of the regression line connecting points 1, 4, 5 and 6 was -1.243 , $r^2=0.979$. *d*, Transmembrane recordings during chaotic activity obtained from a Purkinje cell (upper trace) and a ventricular muscle cell (lower trace) during constant pacing of the Purkinje fibre at a BCL of 137 ms. During chaos conduction from PF to VM was accompanied by foot potentials and

by marked activation delays preceding short-duration VM action potentials. Following a critical activation delay, the VM action potential caused retrograde re-excitation of PF, which in turn, caused a second activation of VM, so that reflection of activation occurred between PF and VM (filled arrow). Reflection also seemed to occur within VM (inferred from biphasic VM action potential; unfilled arrow). Timescale is 100 ms per division.

METHODS. Hearts were removed from adult mongrel dogs of either sex following anaesthesia using pentobarbital sodium (35 mg kg^{-1}). Preparations ($n=8$) of free-running false tendons attached to a 2–3 wide crescent of papillary muscle were excised from the left ventricle and mounted in a tissue bath, as described previously²⁷. Pacing stimuli were delivered to the Purkinje fibre. Following equilibration in normal Tyrode solution, the preparation was superfused with Tyrode solution containing 6.0 mM KCl and 0.5 mM heptanol. Other methods described in the legend to Fig. 1.

absence of reflected responses. Further studies (in progress) may clarify whether 'overlap' of the requirements for chaos and reflection could explain the experimental observations of period two, but not chaos, preceding more (spatially) disorganized activity such as ventricular fibrillation²⁹.

The mechanism we propose for low dimensional chaos in non-pacemaker cardiac cells incorporates general properties common to other nonlinear dynamical systems, ranging from lasers to electronic devices³⁰. Our studies indicate that pharmacological manipulations directed towards the steeply sloped and minimum functions of the model may prevent chaos. Accordingly, drugs that linearize the restitution of action potential amplitude and duration, prevent premature activation and its associated activation delays, or minimize activation delay directly, may suppress the development of cardiac dysrhythmias^{22,31,32}. □

Received 31 July; accepted 14 December 1989.

1. Glass, L., Shrier, A. & Belair, J. in *Chaos* (ed. Holden, A.) 237–256 (Princeton University Press, 1986).
2. Guevara, M. R., Glass, L. & Shrier, A. *Science* **214**, 1350–1353 (1981).
3. Smith, J. M. & Cohen, R. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 233–237 (1984).
4. Glass, L., Goldberger, A. L., Courtemanche, M. & Shrier, A. *Proc. R. Soc. A* **413**, 9–26 (1987).
5. Goldberger, A. L. & Rigney, D. R. in *Dynamic Patterns in Complex Systems* (eds Kelso, J. A. S., Schlesinger, M. F. & Mandell, A. J.) 248–264 (World Scientific, Singapore, 1988).
6. Glass, L., Guevara, M. R., Shrier, A. & Perez, R. *Physica* **70**, 89–101 (1983).
7. El-Sherif, N., Mehra, R., Gough, W. B. & Zeiler, R. H. *Circ. Res.* **51**, 152–166 (1982).
8. Kramer, J. B., Saffitz, J. E., Witkowski, F. X. & Corr, P. B. *Circ. Res.* **56**, 736–754 (1985).
9. Janse, M. J. in *The Heart and Cardiovascular System* (eds Fozzard, H. A. et al.) 1203–1238 (Raven, New York, 1986).
10. Matsumoto, G. et al. *Phys. Lett. A* **123**, 162–166 (1987).
11. Chialvo, D. R. & Jalife, J. *Nature* **330**, 749–752 (1987).
12. Chialvo, D. R. & Jalife, J. in *Cardiac Electrophysiology: From Cell to Bedside* (eds Zipes, D. P. & Jalife, J.) (Saunders, Philadelphia, in the press).
13. May, R. M. *Nature* **261**, 459–467 (1976).
14. Feigenbaum, M. J. *J. Stat. Phys.* **19**, 25–52 (1978).
15. Linsay, P. S. *Phys. Rev. Lett.* **47**, 1349–1352 (1981).
16. Simoyi, R. H., Wolf, A. & Swinney, H. L. *Phys. Rev. Lett.* **49**, 245–248 (1982).
17. Keener, J. P. *J. math. Biol.* **12**, 215–225 (1981).

18. Guevara, M. R., Ward, G., Shrier, A. & Glass, L. in *Computers in Cardiology* 167-170 (IEEE Computer Soc., Long Beach, California, 1984).
19. Guevara, M. R., Shrier, A. & Glass, L. *Am. J. Physiol.* **254**, H1-H10 (1988).
20. Shaw, R. *The Dripping Faucet as a Model Chaotic System* 1-113 (Aerial, Santa Cruz, 1984).
21. Elharrar, V. & Surawicz, B., *Am. J. Physiol.* **244**, H782-H792 (1983).
22. Boyett, M. R. & Jewell, B. R. *Prog. Biophys. molec. Biol.* **36**, 1-52 (1980).
23. Spear, J. F. & Moore, E. N. *Circ. Res.* **35**, 782-792 (1974).
24. Glass, L. & Mackey, M. *From Clocks to Chaos. The Rhythms of Life* 1-248 (Princeton University Press, 1988).
25. Berge, P., Pomeau, Y. & Vidal, C. in *Order Within Chaos. Towards a Deterministic Approach to Turbulence* 1-329 (Wiley, New York, 1984).
26. Gilmour, R. F., Jr in *Cardiac Electrophysiology: From Cell to Bedside* (eds Zipes, D. P. & Jalife, J.) (Saunders, Philadelphia, in the press).
27. Gilmour, R. F., Jr, Davis, J. R. & Zipes, D. P. *Circulation* **76**, 1388-1396 (1987).
28. Antzelevitch, C., Jalife, J. & Moe, G. K. *Circulation* **61**, 182-191 (1980).
29. Ritzberg, A., Adam, D. & Cohen, R. *Nature* **307**, 159-161 (1984).
30. Moon, F. C. *Chaotic Vibrations. An Introduction for Applied Scientists and Engineers* 1-309 (Wiley, New York, 1987).
31. Varro, A., Elharrar, V. & Surawicz, B. *J. Pharmacol. Exp. Ther.* **233**, 304-311 (1985).
32. Colatsky, T. J. & Hogan, P. M. *Circ. Res.* **46**, 543-552 (1980).

ACKNOWLEDGEMENTS. We thank D. C. Michaels and M. Guevara for discussion and M. Buddle and W. Coombs for assisting with some of the experiments. These studies were supported by the NIH.

Expression pattern of the mouse *T* gene and its role in mesoderm formation

David G. Wilkinson, Sangita Bhatt
& Bernhard G. Herrmann*

Laboratory of Eukaryotic Molecular Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

FORMATION of mesoderm is a crucial event in vertebrate development, establishing many of the important features of the body¹⁻⁴. Recent studies have implicated molecules that are similar to growth factors in mesoderm formation in *Xenopus*⁵⁻⁷, but other gene products involved in this process have yet to be identified. Genetic evidence indicates that in the mouse the *T* gene (Brachyury) has a role in the formation and organization of mesoderm. Mice homozygous for mutant alleles of the *T* gene do not generate enough mesoderm⁸, and show severe disruption in morphogenesis of mesoderm-derived structures, in particular the notochord⁹⁻¹². The cloning of the *T* gene¹³ has now allowed us to examine its expression pattern. We report that *T*-gene expression occurs in both early stage mesoderm and its epithelial progenitor, and then becomes restricted to the notochord. This expression pattern correlates with the tissues affected in the *T*-gene mutant, and indicates that the *T* gene has a direct role in the early events of mesoderm formation and in the morphogenesis of the notochord.

We initially examined expression of the *T* gene by RNA blot analysis of RNA from embryos and adult tissues. A 2.1-kilobase (kb) transcript is detected in 8.5- and 9.5-day gastrulation-stage embryos, and at much lower levels in 10.5 to 12.5-day embryos (Fig. 1). Expression of the *T* gene is not detected at later stages of development or in any of the adult tissues that we tested. These data indicate that the period of high-level expression of the *T* gene coincides with the time of mesoderm formation. To examine this further, we analysed the spatial pattern of *T*-gene expression by *in situ* hybridization.

T-gene expression is first detected in early gastrulation-stage embryos. At 7 days of development, transcripts are found in mesoderm cells next to the primitive streak, but not in more lateral mesoderm (Fig. 2*a-d*). In addition, transcripts are found in primitive ectoderm next to the primitive streak (Fig. 2*a-d*), which consists of epithelial cells destined to delaminate and form mesoderm and definitive endoderm¹⁴⁻¹⁶. A similar *T*-gene expression pattern in early mesoderm and primitive ectoderm next to the primitive streak is found in 8.5-day (Fig. 2*e-l*) and

9.5-day (data not shown) embryos. Increasingly anterior regions of these embryos contain progressively more mature mesodermal derivatives, and we examined these in serial longitudinal and transverse sections. *T*-gene expression is down-regulated to undetectable levels in paraxial mesoderm (destined to form somites; Fig. 2*i, j*) and lateral mesoderm (data not shown), and is not detected in the allantois or other extra-embryonic mesodermal tissues (Fig. 2*e-l*). But expression of the *T* gene persists in the notochordal plate of 8.5-day embryos (Fig. 2*e-h, k, l*) and in the definitive notochord of 9.5-day embryos (Fig. 2*m, n*). After gastrulation ends, the only site of *T*-gene expression detected is the notochord, around which cartilaginous cells condense and differentiate to form the vertebral column. Notochordal cells degenerate within each vertebra, but persist to form mucosal cells in the intervertebral discs. Therefore at 17.5 days of development (Fig. 2*o, p*), transcripts of the *T* gene occur in repeated clusters of cells along the length of the spinal column. Overall, these *in situ* hybridization data reveal a relatively widespread expression of the *T* gene in the mesoderm of gastrulating embryos, and a subsequent highly restricted expression in the notochord.

In homozygous *T*-gene mutants, mesoderm formation⁸ and migration¹¹ is disrupted, and the notochord fails to form, possibly because the notochordal plate fails to separate from gut endoderm¹¹. As a result of the deficiency in mesoderm, the allantois does not connect to the maternal circulation and the embryo therefore dies¹⁰. Analysis of these mutants has revealed deletions in the *T* gene, demonstrating that they are true loss-of-function mutants¹³. It is striking that the tissues that express the *T* gene are also those that are disrupted in mutants, indicating that the *T* gene has a direct role in the formation of mesoderm and morphogenesis of the notochord. These observations are consistent with proposals^{9,11} that the defects in somites and the neural tube in *T*-gene mutants are secondary effects presumably reflecting the inductive influence of the notochord on these

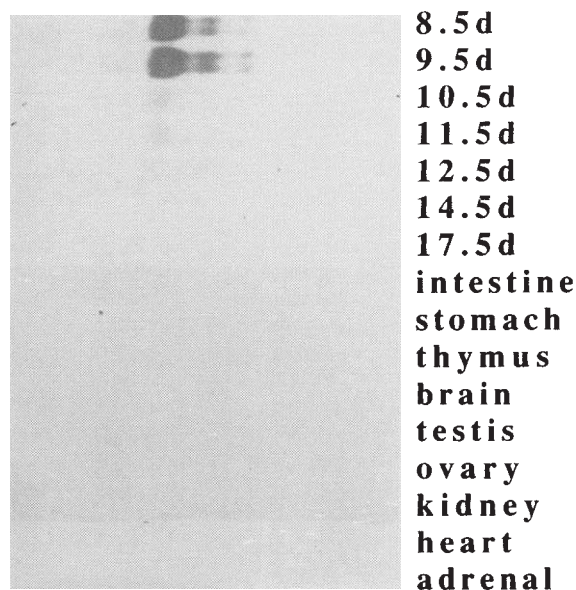


FIG. 1 RNA blot analysis of *T*-gene transcripts in mouse embryos and adult tissues. RNA blots were prepared with total RNA from the indicated mouse embryo and adult tissues, and hybridized with the me75 probe. A principal 2.1-kb transcript is detected in 8.5- and 9.5-day embryos, and at lower levels in 10.5 to 12.5-day embryos.

METHODS. Total RNA was prepared from tissues by the LiCl-urea method²¹ and 20 µg was electrophoresed on a 1% agarose gel containing formaldehyde. Blotting onto nylon membranes and ultraviolet crosslinking was performed by standard methods. The me75 probe¹³ was prepared by random priming²², and hybridization at high stringency was performed as previously described²³. Day, d.

* To whom correspondence should be addressed at: Max Planck Institut für Entwicklungsbiologie, Spemannstrasse 35, 7400 Tübingen, FRG.

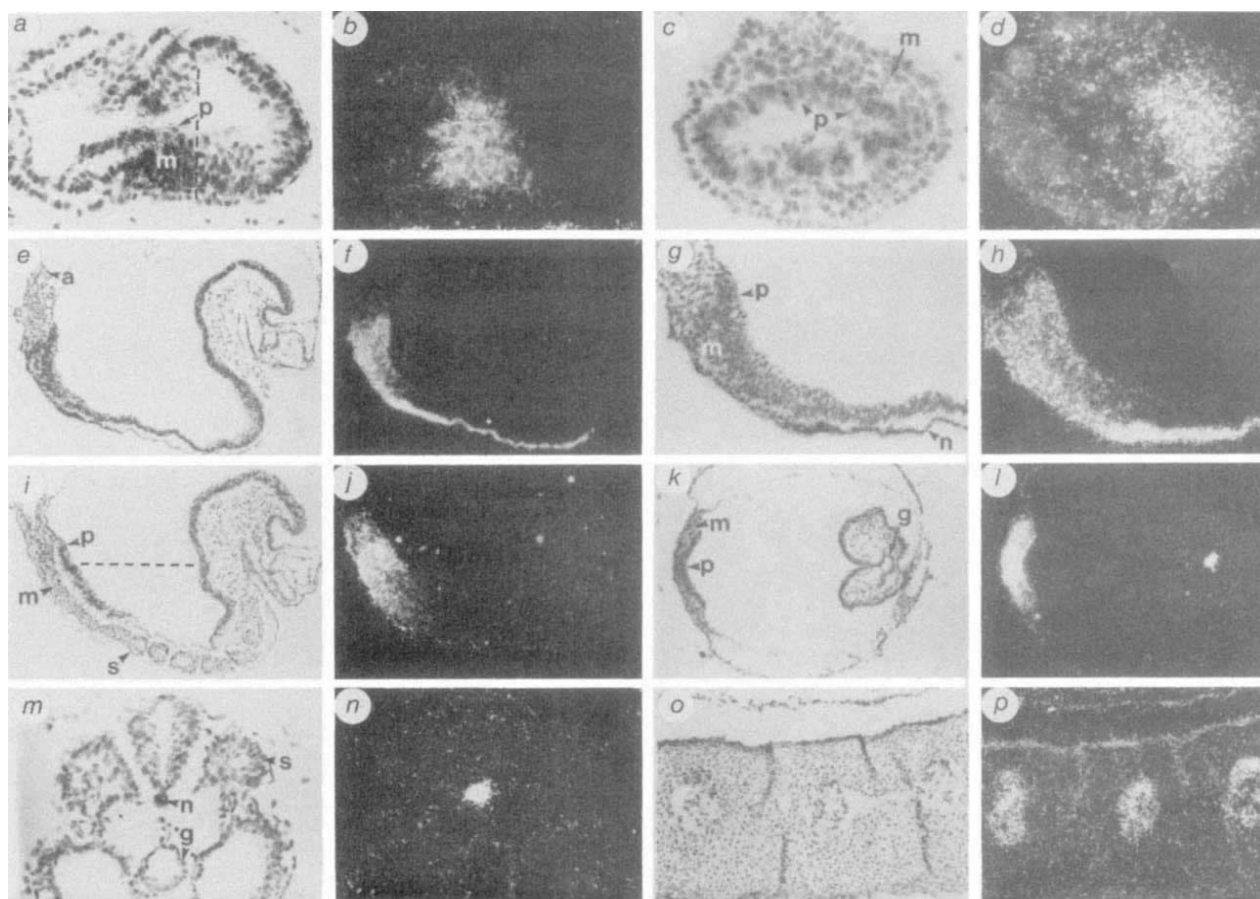


FIG. 2 Analysis of the spatial localization of *T*-gene transcripts in mouse embryos. Sections of mouse embryos were used for *in situ* hybridization analysis with the *T*-gene probe. *a* and *b*, Longitudinal section through 7-day embryo. *c* and *d*, Transverse section through 7-day embryo (the plane of this section is shown by the dashed line in *a*). *e* and *f*, and *i* and *j*, Sagittal and parasagittal longitudinal sections, respectively, of the same 8.5-day embryo. *g* and *h*, Higher magnification view of *e* and *f*. *k* and *l*, Transverse section through an 8.5-day embryo (the plane of this section is shown by the dashed line in *i*). *m* and *n*, Transverse section through spinal region of 9.5-day embryo. *o* and *p*, Longitudinal section through the vertebral column

tissues. Although these secondary effects could be due to the disrupted morphogenesis of the notochord, it is also possible that the *T* gene product has a role in tissue interactions. The observation of fused vertebrae in mice heterozygous for several mutant alleles of the *T* gene^{11,17}, together with the finding reported here of *T*-gene expression in the vertebral discs, indicates a later role for the *T* gene in the morphogenesis and function of notochord-derived tissue.

What cellular process could the *T* gene be involved in? The predicted amino-acid sequence of the *T* gene product indicates that it acts in the cell, possibly in the nucleus, although the possibility that it is secreted cannot be ruled out¹³. It has been suggested that the aberrant morphogenesis in *T*-gene mutants is due to an excessive adhesiveness of mesodermal cells^{11,18}. But if the *T*-gene product is located in the cell, it is not likely to directly mediate cell adhesion. Also, the available data imply that *T*-gene function is not involved solely in mediating or regulating cell adhesion properties; the *T* gene could have a more general role in regulating the cellular changes associated with the formation of mesoderm and the notochord.

Grafting experiments in the mouse embryo have revealed that although the prospective fate of posterior primitive ectoderm is to form mesoderm and endoderm, it is not irreversibly committed to form mesoendoderm before its ingress into the primitive streak¹⁴⁻¹⁶. It should therefore be interesting to ascertain the

of a 17.5-day fetus. *a, c, e, g, i, k, m* and *o*, Bright-field images corresponding to the dark-field photographs of *b, d, f, j, h, l, n* and *p*. Abbreviations: p, primitive ectoderm; m, early mesoderm; a, allantois; s, somites; n, notochord; g, gut. The anterior notochordal plate is intercalated into gut at 8.5 days of development (*k, l*), but has separated from the gut by 9.5 days (*m, n*).

METHODS. *In situ* hybridization analysis using ³⁵S-labelled RNA probes and high-stringency washing was performed as previously described²⁴. The probe sequences correspond to residues from residue 1,760 to the end residue in the 3'-untranslated region of *T*-gene RNA¹³.

timing of mesodermal commitment relative to the onset of *T*-gene expression in the primitive ectoderm. In *Xenopus*, prospective ectodermal cells are induced to migrate and form mesoderm under the influence of growth factors similar to fibroblast growth factor and transforming growth factor- β (refs 5-7). Because the *T*-gene is expressed in both very early mesoderm and its epithelial progenitor in the mouse, it would be pertinent to examine a possible relationship with the expression of growth factors in the primitive streak^{19,20}. A detailed analysis of the *T* gene should provide insight into the molecular mechanisms that underlie mesoderm formation in mammals. □

Received 30 November 1989; accepted 16 January 1990.

1. Snell, G. D. & Stevens, L. C. in *Biology of the Laboratory Mouse* (ed. Green, E. L.) 205-245 (Dover, New York, 1966).
2. Bellairs, R. *Anat. Embryol.* **174**, 1-14 (1986).
3. Beddington, R. S. P. in *Development in Mammals* Vol. 5 (ed. Johnson, M.) 1-32 (Elsevier, Amsterdam, 1987).
4. Slack, J. M. W. *From Egg to Embryo* (Cambridge University Press, 1983).
5. Slack, J. M. W., Darlington, B. G., Heath, J. K. & Godsave, S. F. *Nature* **326**, 197-200 (1987).
6. Kimelman, D. & Kirschner, M. *Cell* **51**, 869-877 (1987).
7. Smith, J. C., Yaqoob, M. & Symes, K. *Development* **103**, 591-600 (1988).
8. Yanagisawa, K. O., Fujimoto, H. & Urushihara, H. *Dev. Biol.* **87**, 242-248 (1981).
9. Chesley, P. *J. exp. Zool.* **70**, 429-459 (1935).
10. Gluecksohn-Schoenheimer, S. *Proc. natn. Acad. Sci. U.S.A.* **30**, 134-140 (1944).
11. Gruneberg, H. *J. Embryol. exp. Morph.* **6**, 424-443 (1958).
12. Fujimoto, H. & Yanagisawa, K. O. *Differentiation* **25**, 44-47 (1983).
13. Herrmann, B. G., Labelit, S., Poustka, A., King, T. R. & Lehrach, H. *Nature* **343**, 617-622 (1990).

14. Beddington, R. S. P. *J. Embryol. exp. Morph.* **64**, 87–104 (1981).
15. Beddington, R. S. P. *J. Embryol. exp. Morph.* **69**, 265–285 (1982).
16. Tam, P. P. L. & Beddington, R. S. P. *Development* **99**, 109–126 (1987).
17. Searle, A. G. *Genet. Res.* **7**, 86–95 (1966).
18. Yanagisawa, K. O. & Fujimoto, H. *J. Embryol. exp. Morph.* **40**, 277–283 (1977).
19. Wilkinson, D. G., Peters, G., Dickson, C. & McMahon, A. P. *EMBO J.* **7**, 691–695 (1988).
20. Paterno, G. D., Gillespie, L. L., Dixon, M. S., Slack, J. M. W. & Heath, J. K. *Development* **106**, 79–83 (1989).
21. Auffray, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303–314 (1979).
22. Feinberg, A. & Vogelstein, B. *Analyt. Biochem.* **137**, 266–267 (1984).
23. Church, G. M. & Gilbert, W. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995 (1984).
24. Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. *Cell* **50**, 79–88 (1987).

ACKNOWLEDGEMENTS. We thank Jim Smith for comments on the manuscript.

HIV susceptibility conferred to human fibroblasts by cytomegalovirus-induced Fc receptor

Jane A. McKeating, Paul D. Griffiths* & Robin A. Weiss

Chester Beatty Laboratories, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

* Virology Department, Royal Free Hospital School of Medicine, Pond Street, London NW3 2PF, UK

THE main receptor for the human immunodeficiency viruses type 1 and 2 (HIV-1 and HIV-2) on T and B lymphocytes, monocytes and macrophages is the CD4 antigen^{1–3}. Infection of these cells is blocked by monoclonal antibodies to CD4^{1,2} and by recombinant soluble CD4^{4–9}. Expression of transfected CD4 on the surface of HeLa and other human cells renders them susceptible to HIV

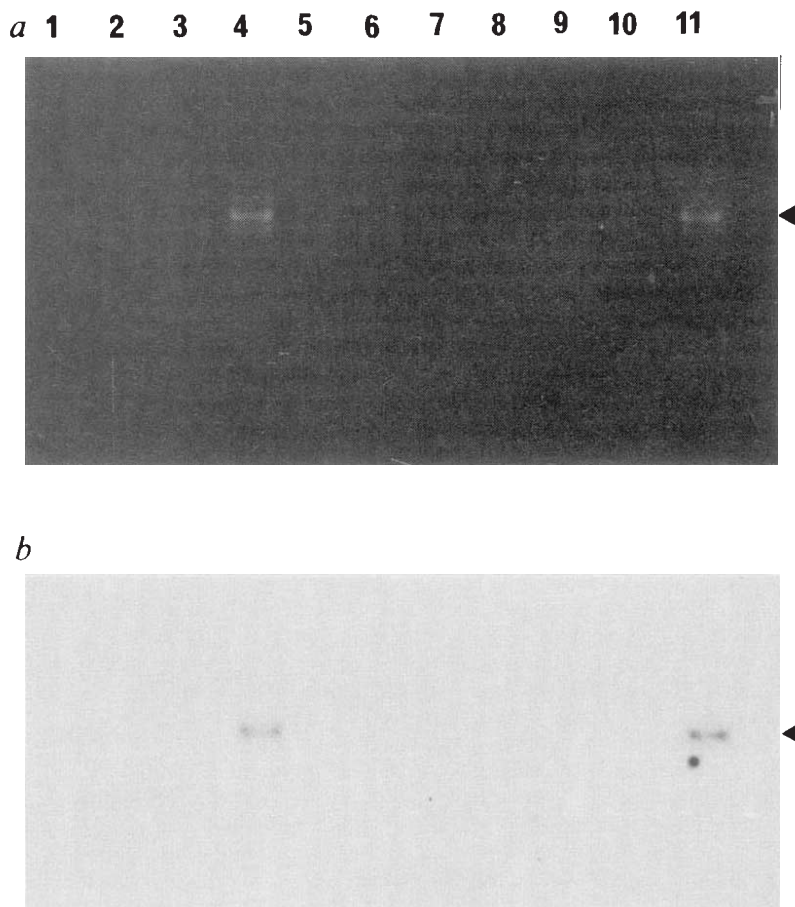
infection¹⁰. HIV–antibody complexes can also infect monocytes and macrophages by means of receptors for the Fc portion of immunoglobulins (FcR)^{11–13}, or complement receptors^{14,15}. The expression of IgG FcRs can be induced in cells infected with human herpes viruses such as herpes simplex virus type 1 (HSV-1)^{16,17} and human cytomegalovirus (CMV)^{18–21}. Here we demonstrate that FcRs induced by CMV allow immune complexes of HIV to infect fibroblasts otherwise not permissive to HIV infection. Infection was inhibited by prior incubation with human IgG, but not by anti-CD4 antibody or by recombinant soluble CD4. Once HIV had entered CMV-infected cells by means of the FcR, its replication could be enhanced by CMV transactivating factors. Synergism between HIV and herpes viruses could also operate *in vivo*, enhancing immunosuppression and permitting the spread of HIV to cells not expressing CD4.

We incubated human embryonic lung (HEL) fibroblasts with HIV at multiplicities of infection (MOI) from 0.01 to 1.0 for T-lymphocytic cell lines. For evidence of HIV infection, we monitored reverse transcriptase activity (RT) and levels of gp120 antigen²² in the medium, and tested for infectious progeny HIV²³. We did not detect HIV infection of fibroblasts within 14 days. We extracted DNA from the cells at 14 days and used the polymerase chain reaction (PCR)²⁴ to test for the presence of a 793-base pair (bp) fragment from the envelope gene (*env*) of HIV. We did not detect any *env* sequences. The HEL fibroblasts were therefore not permissive to HIV infection.

We infected HEL fibroblasts with CMV strain Ad169 (MOI = 0.1) and then used them as targets for HIV infection at the time of maximal FcR expression (5 days). We incubated HIV for 60 min at 37 °C with 1/80, 1/160 and 1/320 dilutions of two human sera. Serum A was anti-HIV positive and anti-CMV negative (HIV⁺/CMV[−])²⁵, with a neutralizing titre for HIV-1 (strain HTLV-IIIB) of 1/160. Serum B was negative for both anti-CMV and anti-HIV antibodies (HIV[−]/CMV[−]). We added

FIG. 1 Superinfection of CMV-infected human lung fibroblasts by HIV–antibody complexes. **a**, Agarose gel (1.5%) of PCR DNA products of HIV-1 (strain HTLV-IIIB)-infected fibroblasts previously infected (tracks 1–4) or uninfected by CMV (tracks 5–8) fibroblasts. HIV was incubated with an HIV[−]/CMV[−] serum (serum B) at a 1/320 dilution (tracks 1 and 5), and an HIV⁺/CMV[−] serum (serum A) at a 1/80 (tracks 2 and 6), 1/160 (tracks 3 and 7) and 1/320 (tracks 4 and 8) dilution. PCR products from uninfected fibroblasts (track 10) and HIV-infected CD4⁺ H9 cells (track 11) were also analysed. **b**, Southern blot of the gel described above, probed with oligonucleotide *env*3.

METHODS. Infected cellular DNA was extracted using standard methods. One microgram of DNA was used in the PCR reaction to amplify a region of the envelope gene (*env*). The primers used were: *env*1, 5'-GTAGTATTGGTAAATGTGA hybridizing to the minus (−) strand at positions 6,467–6,486; *env*2, 5'-CTTAATTGCTATCT hybridizing to the positive (+) strand at positions 7,241–7,260. Amplification reactions were performed in a Techne thermal cycler for 35 cycles as described by Saiki *et al.*²⁴. Primer annealing, extension and denaturation were performed at 94 °C, 47 °C and 72 °C, respectively. The PCR samples were analysed on 1.5% agarose gels for expected band size 793 bp, transferred to Hybond N (Amersham) nylon membranes and hybridized with 5' end-labelled (³²P) oligonucleotide *env*3 5'-CTATAT-CAGTTTATAAAG hybridizing at positions 6,716–6,697. Filters were prehybridized at 60 °C for 1-h in 5 × SSC, 1 × Denhardt's solution, 25% formamide and 10 µg ml^{−1} denatured salmon sperm DNA. A ³²P-labelled heat-denatured probe was added to the prehybridization solution (10⁶ c.p.m. ml^{−1}), and the filter was hybridized overnight. Filters were washed three times with 0.1% SDS, 2 × SSC at 60 °C and twice at room temperature.



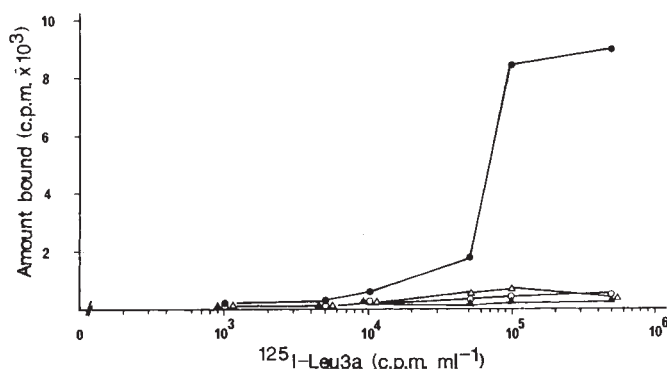


FIG. 2 CD4 cell-surface expression. Binding of ^{125}I -labelled Leu3a to HeLa cells with (●) or without (○) transfected human CD4, or to CMV-infected (▲) or uninfected (△) human lung fibroblasts.

METHODS. Leu3a (Beckton & Dickinson), a monoclonal antibody specific for human CD4, was iodinated using Iodogen (Pierce Chemicals) according to the manufacturer's protocol. All cell lines were seeded at 10^4 cells ml^{-1} in microtitre plates and cultured at 37°C with or without CMV infection. Monolayers were washed three times with HEPES-buffered medium containing 2% BSA. Radiolabelled Leu3a was added at doubling dilutions from 5×10^5 c.p.m. ml^{-1} to 10^3 c.p.m. ml^{-1} for 1 h at 0°C , with four replicates for each dilution. Monolayers were washed as before, and 100 μl NaOH (0.5M) was added to each well for 15 min at room temperature. The resultant cell lysate was counted in a gamma counter.

HIV alone (MOI = 1.0), or HIV pre-incubated with serum A or B, to uninfected and CMV-infected fibroblasts for 60 min at 37°C . We removed unbound virus, and washed the cells twice in serum-free medium. Every 24 h, we tested the culture fluid for RT, gp120 and for infectious HIV and CMV; only CMV-infected fibroblasts were positive in these tests.

Four days after infection, we co-cultivated trypsinized HEL fibroblasts with C8166 cells that are highly permissive for infection by HIV but not CMV. We detected HIV infection in C8166 cells by a syncytial cytopathic effect, immunofluorescence with gp120 monoclonal antibody²⁶, RT and p24 assays. We detected rescued HIV from the CMV-infected fibroblasts treated with HIV incubated with serum A (1/320 dilution). When the syncytial cytopathic effect was visible, HIV-infected C8166 cells had an RT value of 50,200 (control, 3,676), and the medium from these cells contained 7.5 ng p24 per ml.

We used the HEL fibroblasts to prepare DNA for PCR amplification of a 793-bp *env* sequence. Figure 1 shows the analysis of PCR products on an agarose gel. The PCR band of 793-bp in track 4 corresponds to cells infected by HIV complexed with serum A at a dilution of 1/320 (HIV-A). We did not observe PCR bands when uninfected fibroblasts were the targets for HIV infection or when we challenged CMV-infected cells with HIV complexed with serum B (HIV-B). Figure 1b shows a Southern blot of this gel probed with an oligonucleotide internal to the primers used for the PCR amplification step, confirming the identity of the 793-bp *env* band. These results indicate that HIV complexed with low levels of anti-HIV serum can infect fibroblasts previously infected with CMV. This was confirmed with five other HIV-positive antisera diluted beyond their various neutralization titres. The enhancing effect was removed by adsorption with protein A, indicating that it is mediated through immunoglobulin.

To find whether HIV infection of CMV-infected fibroblasts was dependent on CD4, we investigated the expression of CD4 antigen, as well as the effects of Leu3a, a monoclonal antibody specific for human CD4, and recombinant soluble CD4 (sCD4) as agents blocking CD4-dependent infection⁹. The ^{125}I -labelled Leu3a bound to HeLa cells transfected with CD4 (HeLa-CD4

cells), but not to fibroblasts with or without CMV infection (Fig. 2); similar results were obtained by immunofluorescence. Pre-treatment of cells with Leu3a or of virus with sCD4 had no effect on HIV-A infection of CMV-infected fibroblasts, whereas they completely blocked HIV-A infection of HeLa-CD4 cells (Fig. 3).

If the receptor used by HIV-A complexes on CMV-infected fibroblasts is FcR, human serum containing IgG should compete with HIV-A for binding to FcR. We therefore investigated whether excess human serum lacking anti-CMV and anti-HIV antibodies would prevent HIV-A binding to FcR by pre-incubating the target cells with the serum.

Figure 4 shows that 3% and 5% CMV⁻/HIV⁻ human serum blocked HIV-A infection of CMV-infected fibroblasts, without affecting infection of HeLa-CD4 cells. IgG purified from CMV⁻/HIV⁻ human serum had the same effect.

We have demonstrated here that HIV, when complexed with low levels of antibody, can infect CMV-infected fibroblasts. The degree of infection correlated with surface FcR expression induced by CMV, and was maximal at 5 days after infection.

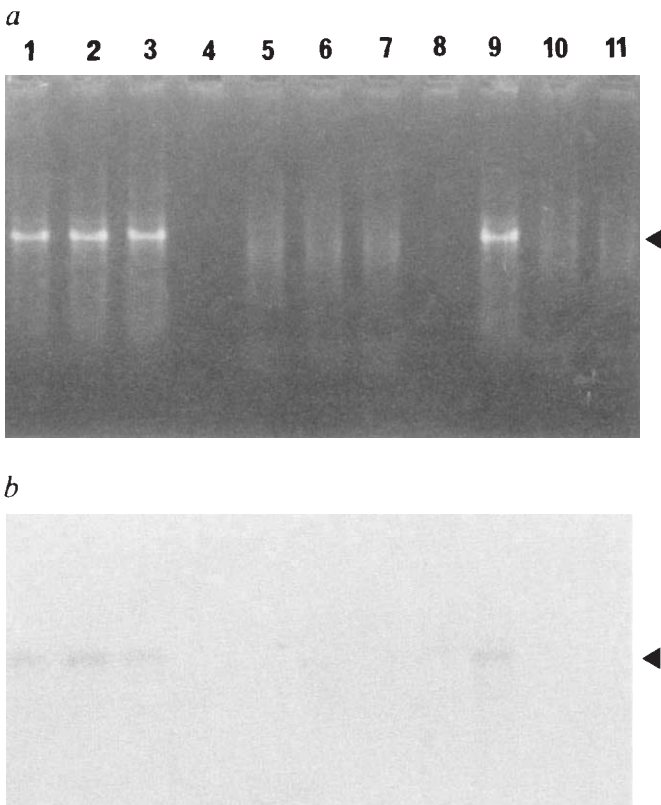


FIG. 3 Effect of Leu3a and sCD4 on HIV-antibody complex infection of CMV-infected fibroblasts. *a*, Agarose gel (1.5%) of PCR products from the following HIV-A-infected cells: fibroblasts (track 5) in the presence of Leu3a (track 6) and sCD4 (track 7); CMV-infected fibroblasts (track 1) in the presence of Leu3a (track 2) and sCD4 (track 3); and HeLa-CD4 cells (track 9) in the presence of Leu3a (track 10) and sCD4 (track 11). *b*, Southern blot of the gel described above was probed with the oligonucleotide *env*3. **METHODS.** HeLa-CD4 cells and fibroblasts either uninfected or CMV-infected (5 days after infection) were treated with saturating amounts of Leu3a ($30 \mu\text{g ml}^{-1}$) at 4°C for 1 h. The cells were then washed and incubated with HIV-A complexes with or without sCD4 at a final concentration of $20 \mu\text{g ml}^{-1}$ for 1 h at 37°C . Four days after infection, DNA was prepared from the cells for PCR amplification as described in the legend to Fig. 1.

Tateno *et al.*²⁷ reported antibody-mediated HIV infection of fibroblastoid cell lines. But we were unable to demonstrate antibody-dependent or antibody-independent HIV infection of the primary HEL fibroblasts that we used.

Similar concentrations of HIV antiserum could neutralize HIV infection of CD4⁺ T-lymphocytes and fibroblasts, indicating that neutralization is independent of CD4 binding. The CD4-independent infection of glial and muscle cell lines is also neutralized by human antiserum to HIV⁹.

The fibroblasts that we used in these experiments did not express CD4 at the cell surface, irrespective of CMV infection. Also, Leu3a and sCD4 had no effect on HIV infection of the fibroblasts, demonstrating that HIV infection of CMV-infected fibroblasts is CD4-independent. Homsy *et al.*¹³ reported that antibody-dependent enhancement of HIV infectivity in peripheral blood monocytes or macrophages is CD4-independent but is blocked by monoclonal antibodies to FcR III.

Three FcRs (I, II and III) have been identified on human leukocytes²⁸. There are no monoclonal antibodies to the CMV-induced FcR, but we exploited its affinity for human IgG²⁹ to

demonstrate enhancement of fibroblast infection by 'opsonized' HIV.

In vivo, CMV shows little tissue tropism, because it is found in most organ systems. Infection by CMV and other herpes viruses inducing the expression of FcR could permit HIV to infect a wide range of cell types. There is a high prevalence of CMV infection in HIV-positive individuals³⁰, with several CMV isolates and multiple sites of infection in patients with AIDS³¹⁻³³. Once HIV-antibody complexes enter CMV-infected cells by means of the induced FcR, HIV replication could be enhanced, because the major immediate early gene of CMV transactivates the long terminal repeat of HIV³⁴. Conversely, HIV superinfection of T cells abortively infected with CMV can lead to productive CMV infection³⁵, suggesting a synergism between these two viruses. Webster *et al.*³⁶ reported that CMV infection correlates with an accelerated disease progression in HIV-infected haemophiliacs. Increased CMV viraemia also occurs during progression from AIDS-related complex to AIDS³⁷. Double infection of CMV and HIV could also result in the release of 'pseudotype' HIV particles bearing CMV envelope glycoproteins, since phenotypic mixing between CMV and retroviruses occurs³⁸. This could lead to the infection of further CD4-negative cells by HIV.

Although antibody-mediated HIV infection of monocytes, macrophages and fibroblasts can occur *in vitro*, the significance of these observations to natural infection and pathogenesis is not known. Current vaccine failure in chimpanzees could reflect an *in vivo* role for enhancing antibodies. Epitopes responsible for inducing enhancing antibodies need to be identified for their exclusion from potential vaccines—virus-enhancing and virus-neutralizing activities have been separated for other virus infections^{39,40}. Our studies show that infection by HIV-antibody complexes mediated by FcR extends to nonhaemopoietic cells and is independent of CD4. □

Received 8 September; accepted 13 December 1989.

1. Dalgleish, A. G. *et al.* *Nature* **312**, 763-767 (1984).
2. Klatzmann, D. *et al.* *Nature* **312**, 767-768 (1984).
3. Sattentau, Q. J. & Weiss, R. A. *Cell* **52**, 631-633 (1988).
4. Smith, D. H. *et al.* *Science* **238**, 1704-1707 (1987).
5. Fisher, R. A. *et al.* *Nature* **331**, 76-78 (1988).
6. Hussey, R. E. *et al.* *Nature* **331**, 78-81 (1988).
7. Deen, K. C. *et al.* *Nature* **331**, 82-84 (1988).
8. Trautner, A., Luke, W. & Karjalainen, K. *Nature* **331**, 84-86 (1988).
9. Clapham, P. R. *et al.* *Nature* **337**, 368-370 (1989).
10. Maddon, P. A. *et al.* *Cell* **47**, 333-348 (1986).
11. Takeda, A., Tuazon, C. U. & Ennis, F. A. *Science* **242**, 580-583 (1988).
12. Homsy, J., Tateno, M. & Levy, J. A. *Lancet* **i**, 1285-1286 (1988).
13. Homsy, J., Meyer, M., Tateno, M., Clarkson, S. & Levy, J. A. *Science* **244**, 1357-1360 (1989).
14. Robinson, W. E., Montefiori, D. C. & Mitchell, W. M. *Lancet* **i**, 790-794 (1988).
15. Robinson, W. E., Montefiori, D. C., Gillespie, D. H. & Mitchell, W. M. *J. AIDS* **2**, 33-42 (1989).
16. Watkins, J. F. *Nature* **202**, 1364-1365 (1964).
17. Westmoreland, D. & Watkins, J. F. *J. gen. Virol.* **24**, 167-178 (1974).
18. Furukawa, T., Hornberger, E., Sakuma, S. & Plotkin, S. A. *J. clin. Microbiol.* **2**, 332-336 (1975).
19. Keller, R., Peitchel, R., Goldman, J. N. & Goldman, M. *J. Immun.* **116**, 772-777 (1976).
20. Westmoreland, D., St. Jeor, S. & Rapp, F. *J. Immun.* **116**, 1566-1570 (1976).
21. Frey, J. & Einsfelder, B. *Eur. J. Biochem.* **138**, 213-216 (1984).
22. Moore, J. & Jarrett, R. *AIDS* **4**, 369-379 (1988).
23. Clapham, P. R. *et al.* *Virology* **158**, 44-51 (1987).
24. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. *Nature* **324**, 163-166 (1986).
25. Berry, N. J., Grundy, J. E. & Griffiths, P. D. *J. virol. Meth.* **13**, 343-350 (1986).
26. Kinney-Thomas, E. *et al.* *AIDS* **2**, 25-29 (1988).
27. Tateno, M., Gonzales-Scarano, F. & Levy, J. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4287-4290 (1989).
28. Unkeless, J. C., Scigliano, E. & Freedman, V. H. A. *Rev. Immun.* **6**, 251-281 (1988).
29. Murayama, T., Natsume-Sakai, S., Shimokawa, K. & Furukawa, T. *J. gen. Virol.* **67**, 1475-1478 (1986).
30. Mintz, L., Drew, W. L., Miner, R. C. & Bruff, E. H. *Ann. intern. Med.* **99**, 326-329 (1983).
31. Reichert, C. M., O'Leary, T. J., Levens, D. L., Simrell, C. R. & Macher, A. M. *Am. J. Path.* **112**, 357-382 (1983).
32. Drew, W. L., Sweet, E. S., Miner, R. C. & Mocarski, E. S. *J. infect. Dis.* **150**, 952-953 (1984).
33. Spector, S. A., Hirata, K. K. & Neuman, T. R. *J. infect. Dis.* **150**, 953-956 (1984).
34. Davis, M. G., Kenney, S. C., Kamaine, J., Pagano, J. S. & Huang, E.-S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8642-8646 (1987).
35. Skolnik, P. R., Kosloff, B. R. & Hirsch, M. S. *J. infect. Dis.* **157**, 508-514 (1988).
36. Webster, A. *et al.* *Lancet* **ii**, 63-65 (1989).
37. Fiala, M., Cone, L. A., Chang, C.-M. & Mocarski, E. S. *AIDS Res.* **2**, 175-181 (1986).
38. Gonczol, E. & Schnitzer, T. J. *J. gen. Virol.* **43**, 691-695 (1979).
39. Cecilia, D., Gadkari, D. A., Kedarnath, N. & Gnosh, S. N. *J. gen. Virol.* **69**, 2741-2747 (1988).
40. Halstead, S. B. *Science* **239**, 476-481 (1988).

ACKNOWLEDGEMENTS We thank D. Evans (Reading University) for oligonucleotide synthesis, R. W. Sweet (Smith Kline Beechams) for sCD4, J. Moore for performing gp120 assays, and M. Collins, M. Marsh and J. Moore for critically reading the manuscript. This study was supported by the AIDS Directed Programme of the MRC.

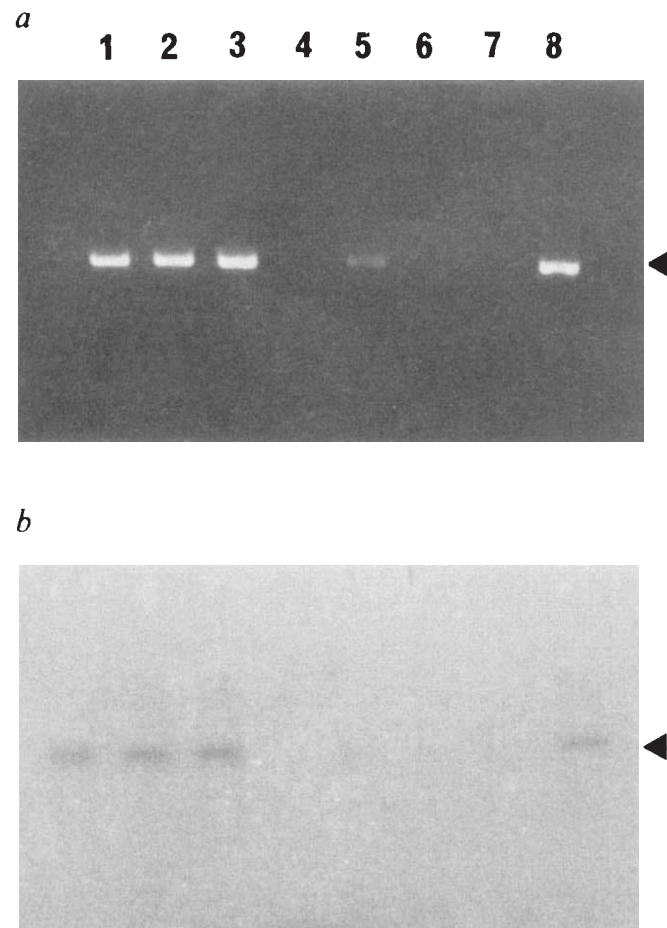


FIG. 4 Effect of blocking FcR by addition of human serum on HIV infection of CMV-infected fibroblasts. *a*, Agarose gel (1.5%) of PCR products from the following HIV-A-infected cells: HeLa-CD4 cells previously incubated with 1% (track 1), 3% (track 2) or 5% (track 3) human serum (HIV-/CMV-); CMV-infected fibroblasts incubated with 1% (track 5), 3% (track 6) or 5% (track 7) of the same serum. Track 8 contains the PCR-amplified material from HTLV-IIIb-infected H9 cells. *b*, Southern blot of this gel probed with the oligonucleotide *env3*.

METHODS. Cells were incubated with PBS, or with 1%, 3% or 5% human serum at 4 °C for 1 h. Cells were then washed and incubated with HIV-A for 1 h at 37 °C. Unbound virus was removed by washing and the cells were maintained at 37 °C. Four days after infection, DNA was prepared for PCR amplification of the 793-bp *env* band as described in the legend of Fig. 1.

Subcellular fate of the Int-2 oncoprotein is determined by choice of initiation codon

Piers Acland, Mark Dixon*, Gordon Peters† & Clive Dickson

Imperial Cancer Research Fund Laboratories, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

† Imperial Cancer Research Fund Laboratories, St Bartholomew's Hospital, Dominion House, 59 Bartholomew Close, London EC1A 7BE, UK

FIBROBLAST growth factors (FGFs) have been implicated in many aspects of cell growth and differentiation both in normal and neoplastic settings^{1,2}. For example, the mouse *int-2* gene, which encodes an FGF-related product³, is a frequent target of proviral activation in carcinomas induced by mouse mammary tumour virus^{4,5}, but apparently functions at discrete stages of normal embryonic development^{6,7}. Six classes of *int-2* messenger RNA have been identified in embryonic cells, each of which is predicted to encode the same 245-amino-acid protein⁸⁻¹⁰. But all known *int-2* transcripts include sequences upstream of the AUG codon presumed to be the initiation codon. Here we report an additional N-terminally extended *int-2* gene product initiated at an in-frame CUG codon. In COS-1 cells transiently transfected with appropriate *int-2* complementary DNAs, the AUG-initiated product is found predominantly in the secretory pathway, whereas the CUG-initiated form is localized to the nucleus. These data indicate that the Int-2 oncoprotein could influence cellular behaviour by two distinct mechanisms.

Previous analysis of the mouse *int-2* gene product identified four principal species, with relative molecular masses between 27,500 and 31,500 (M_r 27.5 K–31.5 K) that could be accounted for by inefficient cleavage of an N-terminal signal peptide and Asn-linked glycosylation at a single consensus site¹¹. Analogous products were observed by translation of synthetic *int-2* RNA *in vitro* or by transient expression of *int-2* cDNAs in COS-1 cells, using a vector based on the simian virus 40 (SV40) early promoter¹¹. However, the number of products depended on the extent of 5' sequences included in the cDNA. For example, Fig. 1a compares the results of *in vitro* translation of two *int-2* cDNAs that differ only in the sequences preceding the AUG codon presumed to be the initiation codon. Because we performed translation in the absence of microsomes, both yielded the expected primary translation product p28.5 (lanes 2 and 3). By retaining a further 432 nucleotides of the upstream *int-2* sequences (lane 3), we obtained an additional product p31.5, presumably corresponding to an N-terminally extended form of Int-2. Although there are no alternative AUG codons suitably positioned in the 5' sequences, we noted an in-frame CUG codon 87 nucleotides upstream of the predicted site of the start of translation (ref. 8, and Fig. 1b). Given the precedent set by the *c-myc* gene¹², initiation at this CUG codon could account for the extended forms of Int-2.

To investigate this possibility, we constructed mutant cDNAs in which the CUG and AUG codons were altered by single nucleotide substitutions (Fig. 2a), and expressed these in COS-1 cells in an SV40-based vector, pKC4 (Fig. 2b). Because the N-terminally extended product was predicted to co-migrate with the glycosylated form of the AUG-initiated Int-2 protein, we first mutated the consensus glycosylation site to replace asparagine 65 with glutamine 65. We visualized the products expressed from the various cDNAs by immunoblotting, using a polyclonal antiserum against a C-terminal peptide from mouse Int-2¹¹. The plasmid designated KC4.1Q, which includes the normal upstream sequences, generated two main products, p28.5

and p31.5, that presumably correspond to the nonglycosylated forms of AUG- and CUG-initiated Int-2 respectively (Fig. 2b, lane 2). A small amount of p27.5 was also apparent, resulting from removal of the signal peptide. Plasmid KC4.2Q, in which the AUG codon has been mutated to AUC (Fig. 2a), yielded only trace amounts of the species p28.5 and p27.5 (Fig. 2b, lane 3). The p31.5 product was unaffected by this mutation. Con-

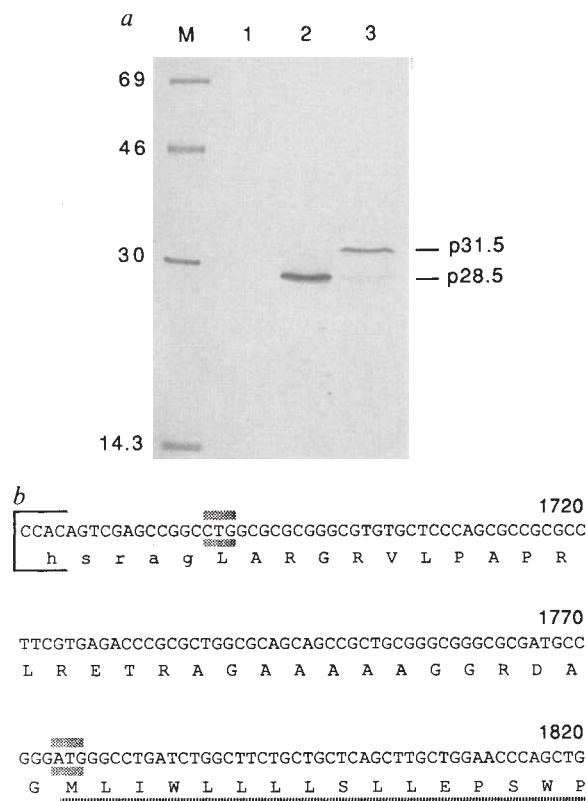


FIG. 1 Translation of cloned *int-2* sequences *in vitro*. **a**, Synthetic RNAs transcribed from mouse *int-2* cDNA sequences using phage SP6 RNA polymerase were used to programme protein synthesis in rabbit reticulocyte lysates in the absence of added microsomes. The ³⁵S-labelled products were treated with RNase to remove band artifacts, and then separated by SDS-PAGE and visualized by autoradiography. Lane 1 shows the background of proteins synthesized in the absence of added RNA. Two cDNA constructs were compared, designated GEM3.2 (lane 2) and GEM4.1 (lane 3), that differed in the lengths of 5' *int-2* sequences preceding the presumed AUG initiation codon. In GEM3.2, only 17 nucleotides were included that has been optimized for initiation according to the consensus proposed by Kozak²⁷ (see ref. 11), whereas in GEM4.1, 432 nucleotides of unaltered leader were present. The primary translation product p28.5 was obtained with GEM3.2 as expected. With the extended cDNA, an additional product, p31.5, was obtained, presumably representing an N-terminal extension of p28.5. The sizes of the various products were calculated relative to ¹⁴C-labelled protein standards (Amersham; lane M) whose M_r ($\times 10^{-3}$) are indicated on the left. **b**, The sequence of mouse *int-2* cDNA is shown beginning at the 5' end of exon 1b, indicated by the open box, and extending to nucleotide 1,820 (refs 8, 9). Amino acids encoded in the presumed reading frame are shown in single-letter code. Those in lower case precede the CTG triplet that could specify an alternative initiation codon (shaded box) while those underlined constitute the signal peptide predicted for Int-2 protein, assuming ATG-directed initiation (shaded box).

METHODS. Construct GEM3.2 was constructed by transferring a *Bgl*II–*Eco*RI fragment from the KC3.2 plasmid¹¹ into the *Bam*HI–*Eco*RI sites in plasmid pGEM3 (Promega). To reconstitute the upstream sequences, a *Sma*I (nucleotide 1,834 in exon 1b) to *Eco*RI (nucleotide 6,268 in exon II) fragment from KC3.2 was transferred into plasmid pGEM4 together with the *Sma*I fragment (nucleotides 1,342–1,834) from *int-2* genomic DNA^{8,9}. Plasmid DNAs were made linear with *Eco*RI and transcribed *in vitro* with SP6 RNA polymerase essentially as recommended by the supplier (Promega). Aliquots of RNA were then translated in rabbit reticulocyte lysates (Promega) in the presence of [³⁵S]methionine and 2 μ l each reaction was analysed by SDS-PAGE in a 12.5% gel¹¹.

* Present address: Plant Biology Laboratory, Salk Institute, La Jolla, California 92037, USA.

versely, replacing the CUG codon with UUG (as in KC4.3Q) drastically reduced the expression of p31.5 but did not affect that of p28.5 (Fig. 2b, lane 4). The double mutant, KC4.4Q, in which both the CUG and AUG were altered, gave only trace amounts of the *int-2* gene products (Fig. 2b, lane 5). This persistence of initiation from the mutated codons is consistent with previous reports that AUC and UUG codons in a favourable context can function as weak initiators for translation in primate cells¹³. It is not clear whether these codons, and the CUG codon, are translated as methionine. Nevertheless, the results clearly demonstrate that N-terminally extended forms of Int-2 protein can be expressed *in vivo* as a result of initiation at a CUG codon.

A central issue is whether FGFs can be exported from cells, and if so, how^{1,2}. Although the prototypic factors, acidic and basic FGF, lack conventional signal peptides and are primarily cytosolic, two transforming oncogenes have been described, kFGF and FGF5, that encode secreted mitogens with typical signal peptides¹⁴⁻¹⁶. Curiously, the AUG-initiated version of the Int-2 protein seems to have intermediate properties in that it enters the secretory pathway, with the concomitant but relatively inefficient removal of a short 17-amino-acid signal¹¹. However, it has not been detected in the medium from expressing cells. We were therefore interested in determining the fate of the CUG-initiated *int-2* gene product, because the additional 29 residues at the N terminus could compromise the properties of the signal peptide. We achieved this by immunofluorescence of transiently transfected COS-1 cells, using either polyclonal antiserum against the C-terminal peptide, or a monoclonal antibody against an internal epitope of mouse Int-2. Figure 3 shows representative examples obtained using KC4.3 and KC4.2 constructs, encoding glycosylated forms of the AUG- and CUG-initiated products, respectively, and the vector KC4. As expected, we found AUG-initiated Int-2 predominantly in the endoplasmic reticulum and Golgi (Fig. 3a, b). By contrast,

CUG-initiated Int-2 was located extensively in the nucleus (Fig. 3c, d). The vector control gave no detectable immunofluorescence (Fig. 3e, f).

This result was independent of transfection and fixation procedures, was achieved with antisera against different Int-2 epitopes, and could be blocked by pre-incubation with the respective peptides. As further confirmation, we also performed cell fractionation experiments to examine the partition of *int-2* products between nuclear, cytosolic and membranous compartments. Again, we found the CUG-initiated species in the nuclear and membranous fractions, whereas we found most of the AUG-initiated products in the membrane fraction (Fig. 4). Antibodies against SV40 T-antigen¹⁷ and the canine docking protein¹⁸ used on replica immunoblots indicated less than 10% cross-contamination of these fractions (data not shown).

The data presented here indicate that the *int-2* gene has evolved so that its product(s) can have different subcellular fates. Most of the products initiated at the AUG codon enter the secretory pathway, but a substantial proportion of those initiated at the CUG codon are diverted to the nucleus. Whether this diversion is directly controlled by a nuclear localization signal in the extended N terminus, or whether it is an indirect response, due to impaired passage through the endoplasmic reticulum, remains to be determined. Although the additional 29-residue sequence is rich in arginine and proline and has superficial resemblance to sequences proposed to be nuclear localization sequences in yeast, there is no obvious equivalent among the established signals described in other systems¹⁹. Indeed, a better candidate for such a consensus (Pro-Arg-Arg-Arg-Lys) is encoded in the first exon of *int-2*, in a position that is common to both CUG- and AUG-initiated forms. Although a similar sequence is present in human Int-2²⁰, it also retains the capacity to encode a CUG-initiated form in which the additional 26-amino-acid sequence is rich in arginine and proline²⁰.

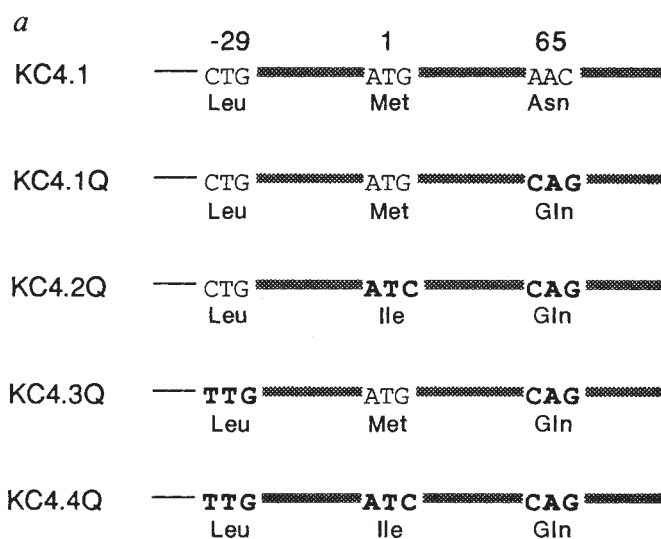
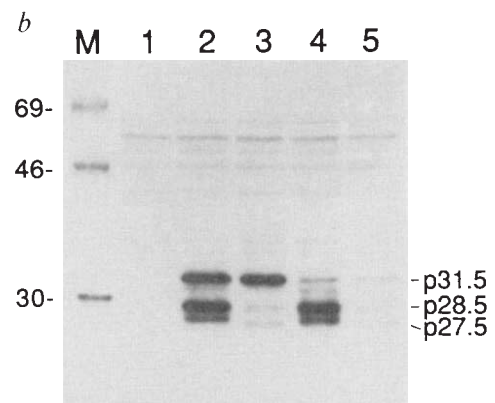


FIG. 2 Expression of mutated *int-2* cDNAs in COS-1 cells. **a**, The *int-2* coding sequences are depicted diagrammatically as shaded lines in which relevant nucleotide triplets corresponding to codons -29, 1 and 65 are indicated. Position 1 refers to the presumed initiator methionine codon. The amino acids encoded by the respective codons are shown underneath. Triplets shown in bold represent mutations relative to the *int-2* cDNA in KC4.1. The mutants designated Q no longer have the consensus site for glycosylation at Asn65; otherwise, they are randomly numbered. The prefix KC4 refers to the SV40 based vector. **b**, Proteins expressed in COS-1 cells transfected with the different cDNAs shown in **a** were analysed by SDS-PAGE and immunoblotting, using a polyclonal antiserum to the C terminus of mouse Int-2. Lane 1 shows results obtained with the KC4 vector alone, whereas subsequent lanes correspond to proteins encoded by KC4.1Q (lane 2),



KC4.2Q (lane 3), KC4.3Q (lane 4) and KC4.4Q (lane 5). Track M shows the ¹⁴C-labelled protein standards. The sizes ($M_r \times 10^{-3}$) and positions of the principal Int-2-specific products are indicated on the right.

METHODS. The *Sma*I fragment encompassing the translation initiation region of *int-2* (see Fig. 1) was transferred from the construct GEM4.1 into the vector M13 and subjected to oligonucleotide-directed mutagenesis using a kit supplied by Amersham. The glycosylation site mutation was introduced in a separate fragment and reinserted into plasmid pGEM4.1 as a *Sma*I-*Kpn*I fragment, generating the construct designated 4.1Q. The mutated *Sma*I fragments were then reassembled in equivalent positions in GEM4.1Q, and each mutation and junction confirmed by DNA sequencing. Finally, the various cDNAs constructed in plasmid pGEM4 were transferred into the KC4 vector, based on the SV40 early region (from D. Hanahan), to generate the plasmids depicted in **a**. The purified plasmid DNAs were introduced into COS-1 cells by electroporation and total cell lysates were prepared 60 h after transfection. The proteins were separated by SDS-PAGE in 12.5% gel and transferred to nitrocellulose by electroblotting. The Int-2-related proteins were detected with antiserum against the C-terminal peptide and ¹²⁵I-labelled protein A (Amersham) as described by Dixon *et al.*¹¹.

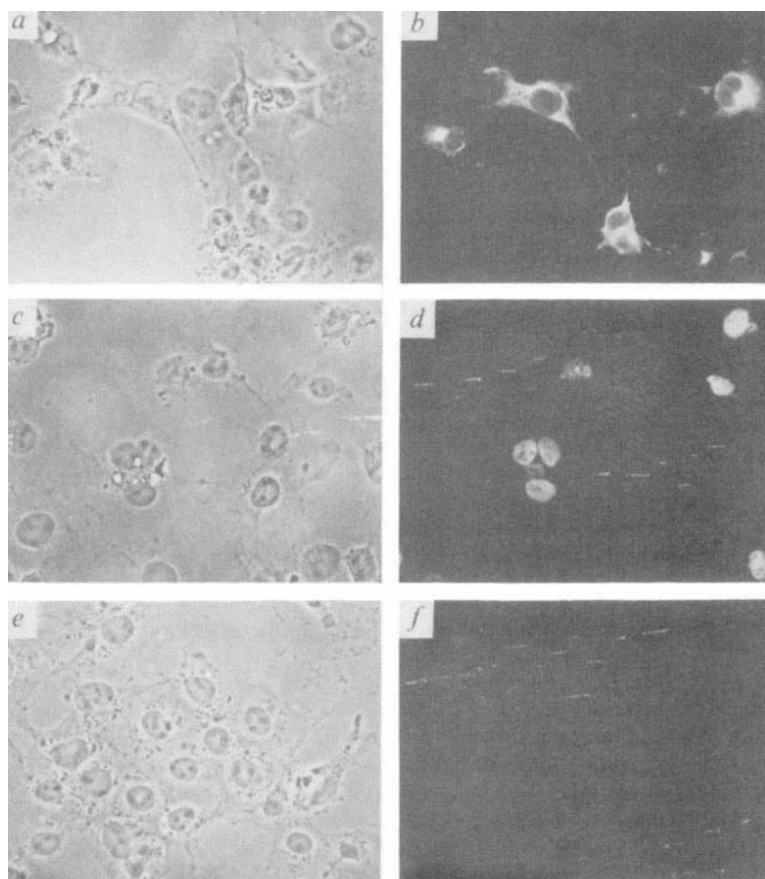


FIG. 3 Subcellular localization of Int-2 proteins by indirect immunofluorescence. The panels show phase contrast (a, c and e) and corresponding fluorescence (b, d and f) micrographs of COS-1 cells transfected with KC4.3 (a and b), KC4.2 (c and d), and KC4 (e and f). KC4.3 and KC4.2 are equivalents of the constructs depicted in Fig. 2a, but without the Q-65 mutation in the glycosylation site.

METHODS. COS-1 cells were transfected with plasmids KC4, KC4.3 and KC4.2 by electroporation, grown on glass coverslips, and 60 h later fixed in 3% paraformaldehyde in PBS and permeabilized in 0.2% Triton X-100 in PBS. After pre-incubation in 1% BSA for 5 min, the coverslips were exposed to a 1:200 dilution of affinity-purified antiserum against a C-terminal Int-2 peptide for 30 min at room temperature. Immune complexes were detected with a rhodamine-conjugated gamma globulin fraction of porcine anti-rabbit serum (Dako Pats) and visualized by fluorescence microscopy using a Zeiss Axiophot microscope.

N-terminally extended forms of human basic FGF that are initiated at any of three alternative CUG codons have recently been described^{21,22}. It will be interesting to determine their subcellular fate, because the additional sequences are similarly rich in arginine and proline. A nuclear location for basic FGF has indeed been reported, but in cells responding to, rather than producing, the factor, and predominantly in the nucleolus²³. There could be an analogous situation for Int-2, because in some of the cells expressing the CUG-initiated product, the staining was distinctly concentrated in the nucleolus. In others, a punctate pattern was observed. The reasons for this variability remain to be determined, but results could be influenced by the

physiological state of individual cells or their position in the cell cycle.

There are precedents for the localization of proteins to both the nucleus and secretory pathway. For example, in cells expressing v-sis, PDGF-related antigens are detectable in the nucleus²⁴, and the proportion of the nuclear forms is increased by mutations in the v-Sis signal peptide²⁵. More significantly, the rat prostatic protein, probasin, shows dual localization as a result of the use of alternative initiation codons²⁶. In contrast to Int-2, the synthesis of a nuclear form of probasin is initiated at an internal AUG codon downstream of the sequence encoding the signal sequence.

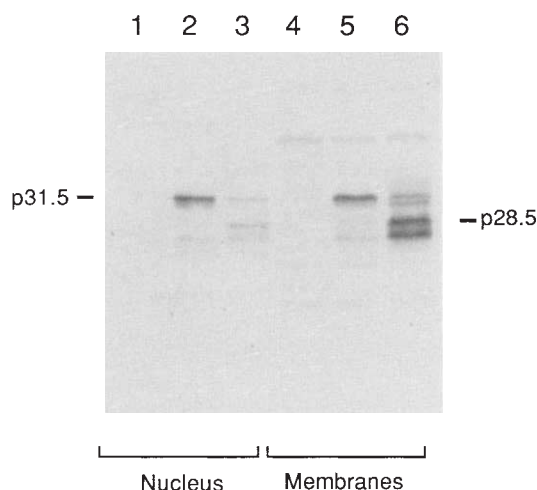


FIG. 4 Subcellular partitioning of Int-2 proteins. Immunoblot of nuclear (lanes 1, 2 and 3) and membrane (lanes 4, 5 and 6) fractions from COS-1 cells transfected with the KC4 vector (lanes 1 and 4), KC4.2Q (lanes 2 and 5) and KC4.3Q (lanes 3 and 6). The main CUG-initiated product from KC4.2Q,

p31.5, is predominantly nuclear, although appreciable amounts are contained in the membrane fraction. The main AUG-initiated products, p28.5 and p27.5, seem to be localized predominantly in the membrane fraction. The smaller products expressed from KC4.2Q are interpreted as deriving from residual initiation at the mutated codon, and a small amount of product resulting from signal-peptide cleavage of p31.5.

METHODS. COS-1 cells transfected with the appropriate plasmids were collected 60 h later and lysed in 5–8 volumes of hypotonic buffer containing 2 mM NaHCO₃, 1 mM MgCl₂, 1 mM EDTA and protease inhibitors. After disruption with a Dounce homogenizer, the cell extract was centrifuged at 800g for 10 min to recover a crude nuclear pellet. The supernatant was centrifuged at 130,000g for 1 h to obtain a mixed mitochondrial and microsomal pellet. The latter was rinsed in HS buffer (10 mM HEPES buffer, pH 7.4, 0.25 M sucrose, 3 mM MgCl₂, and protease inhibitors) containing 0.15 M KCl, and collected by centrifugation. The nuclear pellet was further purified by washing twice in HS buffer, and three times in 10 volumes of 10 mM Tris-HCl buffer, pH 7.4, 1 mM EDTA, 2 mM MgCl₂ and 0.25 M sucrose (TESM). The pellet, recovered by centrifugation at 800g, was resuspended in 40% (w/w) sucrose in 10 mM Tris-HCl buffer, pH 7.4, and 1 mM MgCl₂, and centrifuged through a cushion of 61.5% (w/w) sucrose at 80,000g for 2 h. This purified nuclear pellet was finally washed twice in TSM containing 2% CHAPS to remove residual nuclear envelope, and collected by centrifugation. All manipulations were performed in the cold. Equivalent aliquots of each cell fraction were analysed by SDS-PAGE in a 12.5% gel and Int-2-related proteins were detected by immunoblotting as described previously (Fig. 2, and ref. 11).

The data presented here could indicate a dual function for Int-2 in influencing cell behaviour. The AUG-initiated protein has the hallmarks of an extracellular signalling factor and could be destined for a predominantly paracrine role, whereas the CUG-initiated product is more likely to have an autocrine effect on cells that are expressing *int-2*. Finding whether this effect is positive or negative, or whether it influences the growth or differentiation of the cell, should provide important insights into the contribution of Int-2 during embryogenesis and tumorigenesis. □

Received 29 September; accepted 18 December 1989.

- Gospodarowicz, D., Ferrar, N., Schweigerer, L. & Neufeld, G. *Endocrine Reviews* **8**, 95–114 (1987).
- Rifkin, D. B. & Moscatelli, D. *J. cell. Biol.* **109**, 1–6 (1989).
- Dickson, C. & Peters, G. *Nature* **326**, 833 (1987).
- Dickson, C., Smith, R., Brookes, S. & Peters, G. *Cell* **37**, 529–536 (1984).
- Peters, G. & Dickson, C. in *Cellular and Molecular Biology of Mammary Cancer* eds D. Medina et al. 307–319 (Plenum, New York, 1987).
- Wilkinson, D. G., Peters, G., Dickson, C. & McMahon, A. P. *EMBO J.* **7**, 691–695 (1988).
- Wilkinson, D. G., Bhatt, S. & McMahon, A. P. *Development* **105**, 131–136 (1989).
- Moore, R. et al. *EMBO J.* **5**, 919–924 (1986).
- Smith, R., Peters, G. & Dickson, C. *EMBO J.* **7**, 1013–1022 (1988).
- Mansour, S. L. & Martin, G. R. *EMBO J.* **7**, 2035–2041 (1988).
- Dixon, M. et al. *Molec. cell. Biol.* **9**, 4896–4902 (1989).
- Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. & Eisenman, R. N. *Cell* **52**, 185–195 (1988).
- Peabody, D. S. *J. biol. Chem.* **264**, 5031–5035 (1989).
- Yoshida, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7305–7309 (1987).
- Deili-Bovi, P. et al. *Molec. cell. Biol.* **8**, 2933–2941 (1988).
- Zhan, X., Bates, B., Hu, X. & Goldfarb, M. *Molec. cell. Biol.* **8**, 3487–3495 (1988).
- Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *J. Virol.* **39**, 861–869 (1981).
- Hortsch, M. & Meyer, D. *Eur. J. Biochem.* **150**, 559–564 (1985).
- Dingwall, C. & Laskey, R. A. *Rev. Cell Biol.* **2**, 367–390 (1986).
- Brookes, S., Smith, R., Casey, G., Dickson, C. & Peters, G. *Oncogene* **4**, 429–436 (1989).
- Prats, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1836–1840 (1989).
- Florkiewicz, R. Z. & Sommer, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3978–3981 (1989).
- Bouche, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6770–6774 (1987).
- Yeh, H.-J., Pierce, G. F. & Deuel, T. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2317–2321 (1987).
- Lee, B. A., Maher, D. W., Hannink, M. & Donoghue, D. J. *Molec. cell. Biol.* **7**, 3527–3537 (1987).
- Spence, A. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7843–7847 (1989).
- Kozak, M. *J. Cell Biol.* **108**, 229–241 (1989).

ACKNOWLEDGEMENTS. We thank Francis Fuller-Pace and John Armstrong for discussions, Roger Watson and Gerard Evan for comments on the manuscripts, David MacAllan, John Armstrong and Ed Harlow for anti-peptide Int-2 serum, anti-docking protein serum and anti-T antigen antibodies, respectively, and Iain Goldsmith for the preparation of synthetic oligonucleotides.

Retinoblastoma in transgenic mice

Jolene J. Windle*, Daniel M. Albert†, Joan M. O'Brien†, Dennis M. Marcus†, Christine M. Disteche‡, Rene Bernards§ & Pamela L. Mellon*||

*Regulatory Biology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

†Howe Laboratory of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, Massachusetts 02114, USA

‡Department of Pathology, University of Washington, School of Medicine, Seattle, Washington 98195, USA

§Massachusetts General Hospital Cancer Center, Harvard Medical School, Bldg. 149 13th Street, Charlestown, Massachusetts 02129, USA

RETINOBLASTOMA, a malignancy of the eye occurring in young children, has been widely studied as a model for genetic predisposition to cancer. This disease is caused by mutations in both alleles of an anti-oncogene (the retinoblastoma gene, *Rb*) that inactivate or eliminate the *Rb* encoded protein, p105^{Rb} (refs 1 and 2). Here we report that expression of a viral oncogene, the simian virus 40 T antigen, in the retina of transgenic mice produces heritable ocular tumours with histological, ultrastructural and immunohistochemical features identical to those of human retinoblastoma. Furthermore, we demonstrate a specific association³ between p105^{Rb} and T antigen in mouse retinoblastoma tumour cells. Thus,

the occurrence of these tumours is *in vivo* evidence for oncogenesis due to the ocular-specific expression of an *Rb*-binding oncoprotein that can functionally inactivate the *Rb* protein. As an animal model for heritable retinoblastoma, these mice should allow the study of the ontogeny, pathogenesis and treatment of this malignant disease.

Fifteen lines of transgenic mice were derived from fertilized ova microinjected with a chimaeric gene containing the protein coding region of the simian virus 40 (SV40) T antigen (*Tag*) driven by the promoter of the luteinizing hormone β -subunit gene (*LH β*)⁴. The introduction of *Tag* into transgenic mice has produced tumours of a variety of tissues^{5,6}, both by directed oncogenesis^{7–9} and by ectopic expression¹⁰. The *LH β -Tag* hybrid gene directed a low level of *Tag* expression specifically to gonadotrope cells of the anterior pituitary in most of the lines of mice, but did not result in observable histological abnormalities in this organ until 10–12 months of age. However, a single male founder developed bilateral ocular neoplasms at about 5 months of age. The phenotype was heritable with complete penetrance in transgenic offspring. Retinal tumours were first observed at about 2 months of age, and filled the vitreous cavity by 5–6 months of age.

A variety of tissues from this line of transgenic mice, including ocular tumours, were assayed for *Tag* messenger RNA by northern blotting. *Tag* mRNA was present at marginally detectable levels in the pituitary, and at considerably higher levels in the eye, with the level increasing as the tumours developed (Fig. 1). *Tag* mRNA was absent from other tissues.

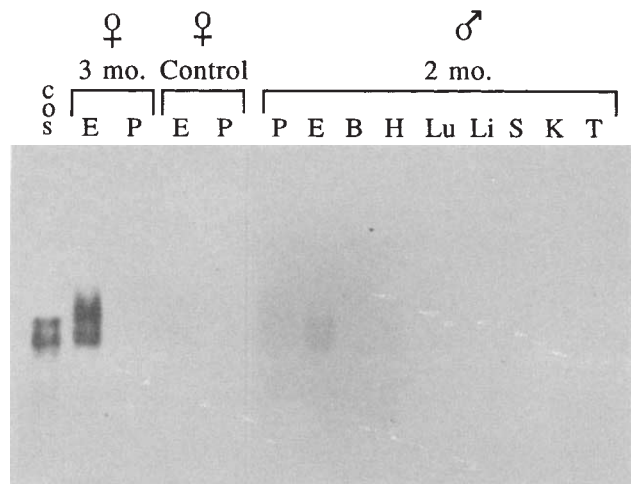
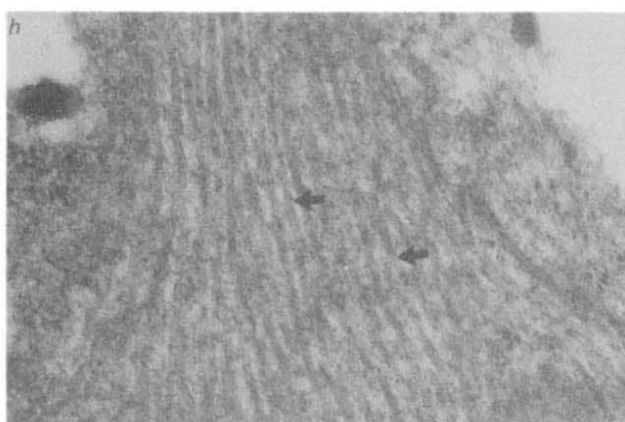
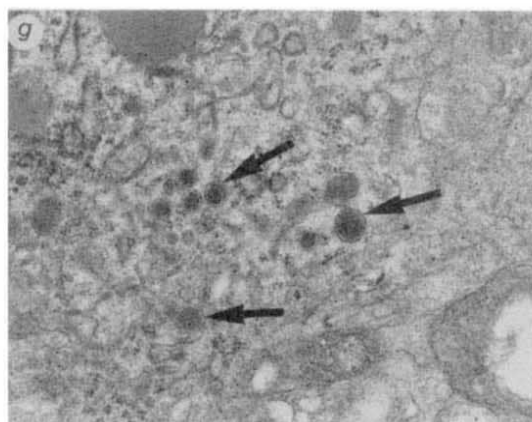
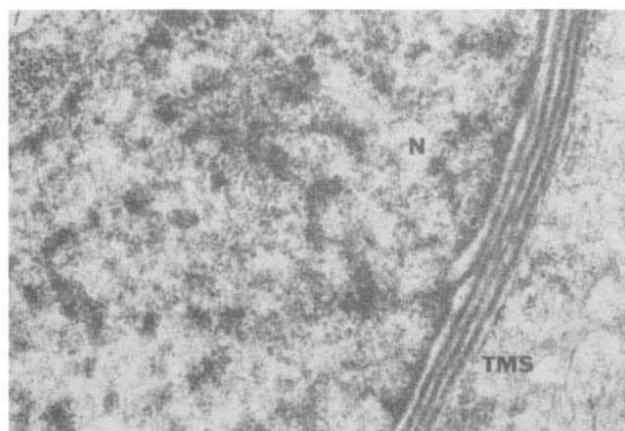
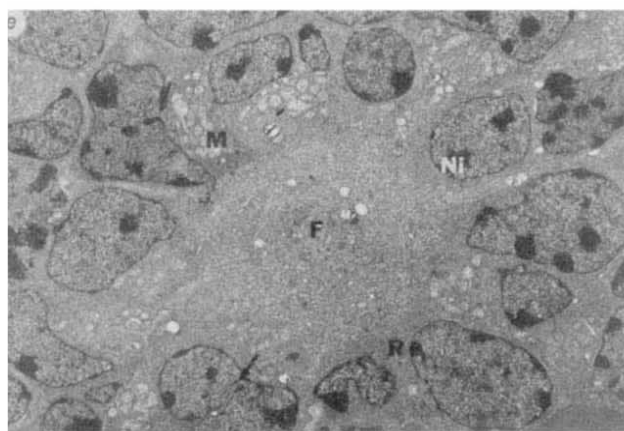
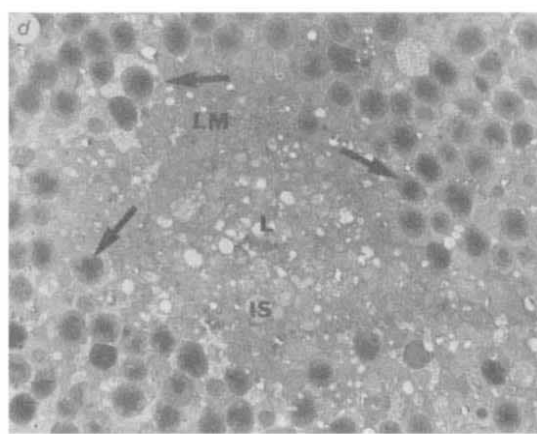
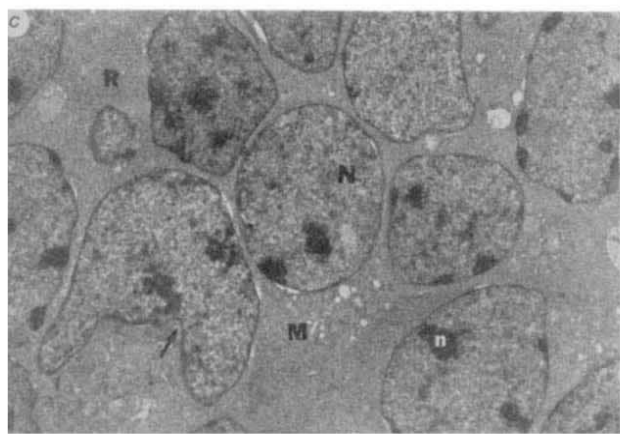
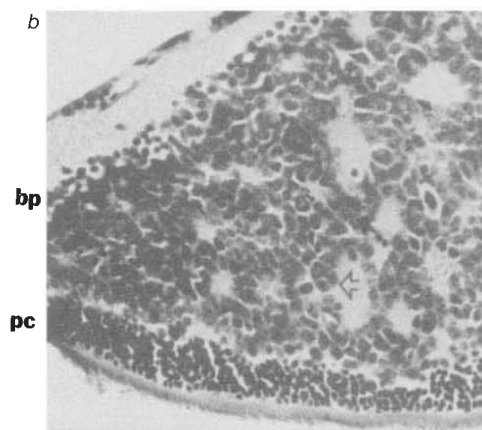
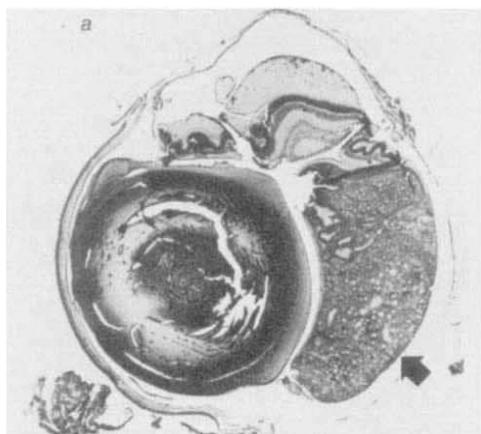


FIG. 1 SV40 *Tag* expression in various tissues of *Tag Rb* mice. Total RNA was prepared as described³⁰ from the eye and pituitary of a 3-month-old transgenic female in which development of ocular tumours was readily apparent, and from various tissues of a 2-month-old transgenic male in which the presence of an ocular tumour was only apparent on histological examination. In addition, RNA was prepared from the eye and pituitary of a nontransgenic control female and from COS cells, which constitutively express *Tag*³¹. *Tag* mRNA levels were assayed by northern analysis³², using a probe specific for *Tag* (SP6TagHB, provided by Y. Gluzman). Total RNA (10 μ g) was present in each lane, except for the pituitary samples, which contained either 5.0 μ g (the transgenic and control females) or 7.6 μ g (the transgenic male). Abbreviations: E, eye; P, pituitary; B, brain; H, heart; Lu, lung; Li, liver; S, spleen; K, kidney; T, testes; mo, months.

METHODS. The *LH β -Tag* transgene contains the human LH β -subunit promoter region⁴ from –1.09 kilobases to +9 base pairs linked to the SV40 early region from the *Bgl*I site to the *Bam*HI site³³. This fragment contains the coding region for *Tag* and small t antigen, including the translation initiation and transcription termination sites, but lacks the SV40 early promoter. The hybrid gene was precisely excised from the plasmid vector, purified and injected into fertilized one-cell embryos essentially as described³⁴. The F₂ embryos were derived from matings of CB6F1/J (C57BL/6J $\times$ BALB/cJ) males and females, obtained from the Jackson Laboratory.

|| To whom correspondence should be addressed.



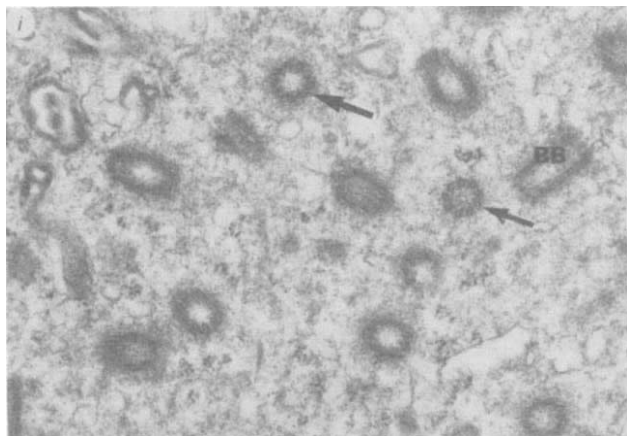


FIG. 2 Light and transmission electron microscopy of ocular tumour sections. Eyes were enucleated at about 5 months of age. The globes were fixed in 5% buffered formalin or 3% phosphate buffered glutaraldehyde-sucrose and processed for light (a, b) or transmission electron microscopy (c-i), respectively. a, Low-power view of eye containing retinoblastoma (arrow). b, Haematoxylin and eosin stain of retinoblastoma cells (arrow) interposed between the bipolar cell layer (BP) and photoreceptor cells (PC). The tumour is composed of both undifferentiated cells and rosettes. Magnification, $\times 350$. c, Area of tumour cells similar in appearance to undifferentiated tumour cells found in human retinoblastoma. The cells have prominent round nuclei (N), multiple nucleoli (n), and in some cases indentations of nuclear membrane (arrow). The cytoplasm is relatively scanty and contains ribosomes (R) and mitochondria (M). Magnification, $\times 9,150$. d, Highly differentiated rosette composed of photoreceptor-like cells (arrows) having similarity to Flexner-Wintersteiner rosette. Delicate limiting membrane (LM) is present. Disorganized inner segment material (IS) fills lumen (L). Magnification, $\times 630$. e, Cuboidal cells forming a rosette configuration around a fibrous matrix (F) resembling a Homer Wright rosette. Nuclei contain multiple nucleoli (Ni), fine marginal chromatin, and occasional indentations of nuclear surface (arrows). Mitochondria (M) and aggregates of ribosomes (R) are present in the cytoplasm. Magnification, $\times 3,400$. f, Nuclear membrane of mouse tumour cells shows triple membrane structure (TMS) with dense chromatin-like material associated with the nuclear membrane. The Figure includes a portion of the nucleus (N). Magnification, $\times 18,000$. g, Neurosecretory granules (arrows) in mouse retinoblastoma cell cytoplasm, measuring 900–1,100 Å in diameter. These granules are frequently seen in human retinoblastoma cells and in the amacrine cells of the retina. Magnification, $\times 22,440$. h, Microtubules (arrows) in the cytoplasm of mouse tumour cells, measuring 200–250 Å in diameter. Magnification, $\times 42,750$. i, Transverse sections of cilia with a pattern of 9+0 microtubules (arrows) in mouse tumour cell cytoplasm. Similar cilia occur in human retinoblastoma cells as well as photoreceptor cells. Basal bodies (BB) are also present. Magnification, $\times 9,920$.

The histology of retinoblastoma is well described^{11,12}, and the features of this malignancy were closely paralleled by the murine tumours (Fig. 2a). Human retinoblastoma is a small cell tumour of the sensory retina. The predominant cell type has a large hyperchromatic nucleus with minimal cytoplasm. These poorly differentiated cells have a high mitotic index and may be precursors of more-differentiated tumour cells displaying features of photoreceptor cells. Perivascular cuffs of viable tumour cells surround retinal vessels and contrast prominently with areas of necrosis and calcification near these vessels¹³. These features were observed in the murine tumours both by light microscopy (Fig. 2b) and by electron microscopy (Fig. 2c).

Human retinoblastoma cells undergo photoreceptor cell differentiation to form characteristic rosettes, first described by Flexner¹⁴ and Wintersteiner¹⁵. These rosettes are composed of cuboidal cells attached at apical ends to terminal bars, which surround a central lumen and are uniquely associated with retinoblastoma. A second structure, the Homer Wright rosette, is composed of a single row of less-differentiated cells arranged in a radial pattern around a tangle of fibrous cytoplasmic processes¹⁶. Rosettes of both types were observed by light and electron

microscopy in every transgenic mouse tumour examined (Flexner-Wintersteiner rosettes are shown in Fig. 2b and d, and Homer Wright rosettes are shown in Fig. 2b and e).

Ultrastructural studies of the murine tumours revealed additional features that are characteristic of human retinoblastoma. Lamelliform nuclear membranes, in which a central dense layer of either granular or fibrillar chromatin is bounded on both sides by membrane, were often observed. Neurosecretory granules, cytoplasmic microtubules, and cilia with a '9+0' pattern of microtubules in which the shafts lack central processes, were also observed (Fig. 2; f, g, h and i, respectively). These ultrastructural features are characteristic of human photoreceptor cells and of both human and murine retinoblastoma cells differentiating along the photoreceptor cell lineage.

Immunohistochemical staining for neuron-associated antigens in tumours from transgenic mice also provided results consistent with those obtained from human retinoblastoma. Kyritsis and colleagues demonstrated¹⁷ both neuron specific enolase (NSE) and glial fibrillary acidic protein (GFAP) in undifferentiated human retinoblastoma cells. With differentiation, retinoblastoma cells showing glial morphology become negative for NSE, whereas those with neural features become negative for GFAP¹⁷. The transgenic murine tumours stained positively for NSE and negatively for S-100 and vimentin. GFAP was identified in the supporting stroma of the tumour but not in the tumour itself. These immunohistochemical features are consistent with a neural pattern of differentiation and are identical to the staining pattern of the well characterized human retinoblastoma cell line Y-79 (ref. 18; data not shown).

In the eye, malignant extension of human retinoblastoma occurs by local invasion or by vitreous seeding^{11,12}. Endophytic tumours grow from the inner retina into the vitreous cavity. Exophytic tumours extend towards the choroid and produce retinal detachment; these tumours can eventually extend along ciliary vessels and nerves to the orbit. Invasion of the optic nerve with malignant extension to the brain also occurs^{11,12}. Endophytic and exophytic growth are most commonly observed in the same eye. Transgenic murine tumours displayed both endophytic and exophytic growth, with invasion of retina and choroid, vitreous seeding and optic nerve invasion. These tumours arose bilaterally, with several foci, as occurs in heritable human retinoblastoma.

The remarkable fidelity with which the retinal tumours in these mice displayed the highly specific and complex ultrastructural features of human retinoblastoma, indicates that there may be a common underlying mechanism for tumorigenesis at the molecular level. One possibility is that the *Rb* locus was disrupted by the transgene or suffered some other mutation, and tumorigenesis resulted from a frequent somatic mutation of the second allele (as is the case in heritable human retinoblastoma). The transgene was inherited with a mendelian pattern (50% of offspring receiving the transgene), indicating a single chromosomal integration site. To establish whether the transgene integrated into one allele of the *Rb* locus, the transgene was mapped on metaphase chromosomes by *in situ* hybridization using the *Tag* coding region as a probe. The mouse *Rb* gene has been mapped to chromosome 14 (refs 19 and 20); but, the *Tag* transgene was located at a single site on chromosome 4 between bands C5 and C7 (Fig. 3). It is also unlikely that the mouse retinoblastoma resulted from any other mutation in the *Rb* locus, because these animals had been bred to wild-type animals through 10 successive generations and the fully penetrant phenotype correlated 100% with the presence of the *Tag* transgene. In addition, because p105^{Rb} is expressed ubiquitously^{1,2}, integration into this gene would not explain the specific expression of the transgene in the eye. The retinal-specific expression was probably a result of integration close to a transcriptional regulatory sequence capable of directing retinal-specific expression, such as an enhancer for a retinal-specific gene.

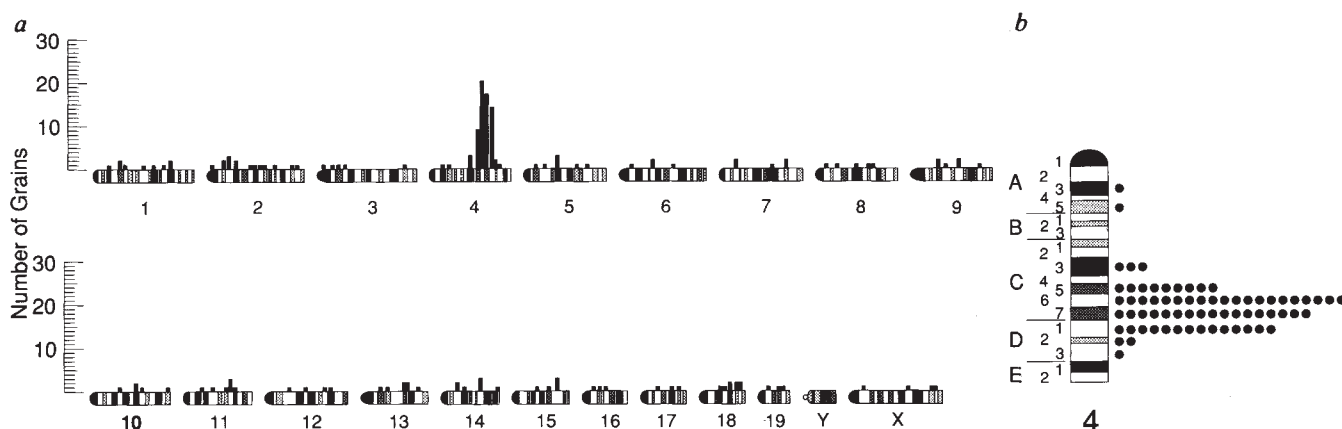


FIG. 3 Regional chromosome localization of the transgene. The plasmid SP6TagHB containing the *Tag* coding region of SV40 was labelled by nick translation using ^3H -labelled nucleotides to a specific activity of 6×10^7 c.p.m. μg^{-1} . *In situ* hybridization to metaphase chromosomes from lymphocytes of a transgenic mouse was carried out using 0.02 μg per μl probe as described³⁵, and exposed for 7 days. A total of 81 metaphase cells were examined and chromosomes were identified by Q-band staining.

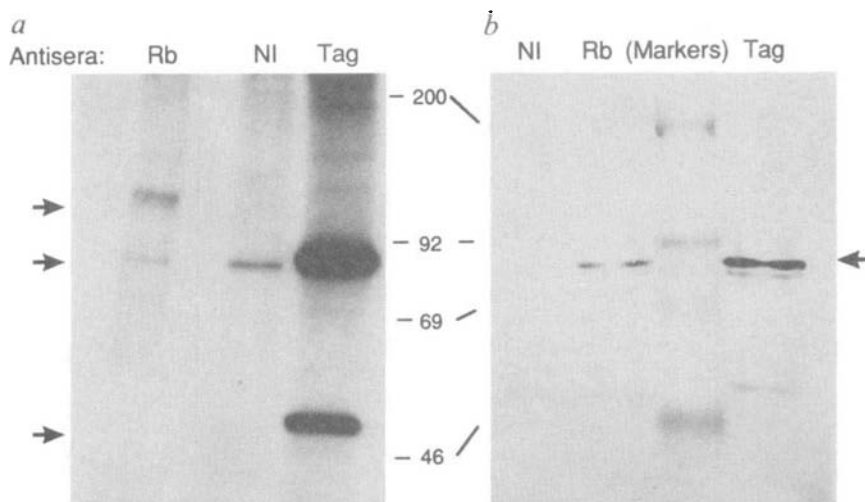
A more intriguing possibility for the ontogeny of tumorigenesis in these mice is indicated by the demonstrations of direct interactions between p105^{Rb} and T antigen³ or other oncoproteins^{21,22} in transformed fibroblasts or *in vitro*^{22,23}. One mechanism by which T antigen could mediate tumorigenesis is through direct inhibition of the function of the Rb protein²⁴. To demonstrate that both p105^{Rb} and Tag proteins were expressed, retinoblastoma cells were metabolically labelled with ^{32}P and protein extracts were immunoprecipitated with three antibodies: anti-mouse p105^{Rb} antiserum (Rb-138; ref. 25), anti-Tag antibody (PAb416; refs 3 and 26), and preimmune serum. Figure 4a demonstrates that the anti-mouse p105^{Rb} antiserum precipitated a protein of relative molecular mass 105,000 (M_r 105 K), whereas the anti-Tag antibody precipitated a 90K protein (TAG) and a 53K protein (presumed to be p53, another cellular oncoprotein that associates with Tag²⁷⁻²⁹). Both p105^{Rb} and Tag were therefore expressed in the murine retinoblastoma

cells. From this experiment, it was difficult to assess whether p105^{Rb} and Tag were associated and could be co-immunoprecipitated, because a protein precipitated by both the preimmune serum and the anti-p105^{Rb} antiserum migrated to about the same position as Tag. In addition, p105^{Rb} would not be expected to be visible in the anti-Tag antibody lane, because Tag binds only the unphosphorylated form of p105^{Rb} (ref. 23), and these cells were labelled with ^{32}P . In order to demonstrate an association between p105^{Rb} and Tag, unlabelled extracts from ocular tumour cells were immunoprecipitated with anti-p105^{Rb} antiserum and analysed by western blotting using the anti-Tag antibody. Tag was co-immunoprecipitated by the anti-p105^{Rb} antiserum, indicating a direct interaction between p105^{Rb} and Tag (Fig. 4b). These observations strongly support a mechanism for the development of retinoblastoma in which the normal function of p105^{Rb} is impaired through binding to Tag specifically expressed in ocular tissues, and provide an

cells. From this experiment, it was difficult to assess whether p105^{Rb} and Tag were associated and could be co-immunoprecipitated, because a protein precipitated by both the preimmune serum and the anti-p105^{Rb} antiserum migrated to about the same position as Tag. In addition, p105^{Rb} would not be expected to be visible in the anti-Tag antibody lane, because Tag binds only the unphosphorylated form of p105^{Rb} (ref. 23), and these cells were labelled with ^{32}P . In order to demonstrate an association between p105^{Rb} and Tag, unlabelled extracts from ocular tumour cells were immunoprecipitated with anti-p105^{Rb} antiserum and analysed by western blotting using the anti-Tag antibody. Tag was co-immunoprecipitated by the anti-p105^{Rb} antiserum, indicating a direct interaction between p105^{Rb} and Tag (Fig. 4b). These observations strongly support a mechanism for the development of retinoblastoma in which the normal function of p105^{Rb} is impaired through binding to Tag specifically expressed in ocular tissues, and provide an

FIG. 4 Immunoprecipitation and western-blot analysis of Tag and p105^{Rb} from retinoblastoma tumour cells. *a*, Extracts from ^{32}P orthophosphate-labelled cultures of transgenic ocular tumours grown in nude mice were immunoprecipitated with a mouse monoclonal antibody to Tag (PAb-416; ref. 26), rabbit preimmune serum (138-NI; ref. 25), and a polyclonal antiserum directed against a portion of the murine p105^{Rb} (Rb-138; ref. 25). *b*, Unlabelled proteins extracted directly from transgenic ocular tumours were immunoprecipitated with the same three antisera and western-blotted with the antibody against Tag.

METHODS. *a*, About 10^7 retinoblastoma cells from ocular transgenic tumours were introduced subcutaneously into nude mice. Nude mouse tumours were transferred to tissue culture and grown in DMEM, 15% horse serum and 5% fetal bovine serum. Retinoblastoma cells grew in suspension in characteristic clusters and any fibroblasts were removed due to adherence to plastic. Uniform populations of nonadherent cells were studied by light and electron microscopy and determined to be a pure population of retinoblastoma cells (data not shown). Cell cultures were treated with phosphate-free DMEM for 2 h at 37 °C and then labelled with 5 mCi of ^{32}P orthophosphate (NEN) in phosphate-free DMEM for 4 h at 37 °C. Cells were lysed in ELB buffer (0.15 M NaCl, 0.1% Nonidet P-40, 50 mM HEPES buffer, 1 mM phenylmethylsulphonyl fluoride, 5 mM EDTA, 0.5 mM dithiothreitol). Supernatants were precleared with protein A Sepharose, then incubated with appropriate antibodies for 8 h on ice. Proteins were immunoprecipitated with protein A Sepharose, washed with ELB buffer, and separated on a 7.5%



SDS-polyacrylamide gel. *b*, Retinoblastoma cells from 10 ocular transgenic mouse tumours were dissected and lysed in ELB buffer. Nuclear proteins were precipitated and electrophoresed as described above, then transferred to nitrocellulose by electrophoresis. The filter was blocked by incubation in TBST (10 mM Tris-HCl buffer, pH 8.0, 150 mM NaCl, 0.05% Tween-20, 1% BSA) then incubated with 15 ml of a 1:100 dilution of PAb-416 for 1 h, washed in TBST and incubated with alkaline phosphatase-conjugated rabbit anti-mouse IgG (Promega). Colorimetric assays were performed according to the manufacturer's specifications.

explanation for the remarkable identity in morphology and malignant characteristics of the transgenic and human retinoblastomas.

Heritable retinoblastoma has greatly advanced our knowledge of the genetics of cancer. However, this malignancy has been observed exclusively in humans, and the absence of a heritable animal model has limited detailed understanding of the pathogenesis of this disease. We have described a line of transgenic mice that develop intraocular neoplasms with the histologic, ultrastructural, immunohistochemical and invasive features of human retinoblastoma, created by specific expression of the *Tag* oncogene in the retina. The demonstration that *Tag* interacts with the *Rb* anti-oncogene product, establishes that this association can occur in retinoblastoma tumours *in vivo* and provides strong evidence for its role in oncogenesis. This model provides an opportunity to study the process of malignant transformation in retinoblastoma *in vivo* from inception through extension and metastases. □

Received 6 November; accepted 13 December 1989.

1. Friend, S. H. *et al.* *Nature* **323**, 643–646 (1986).
2. Lee, W.-H. *et al.* *Science* **235**, 1394–1399 (1987).
3. DeCaprio, J. A. *et al.* *Cell* **54**, 275–283 (1988).
4. Talmadge, K., Vamvakopoulos, N. C. & Fiddes, J. C. *Nature* **307**, 37–40 (1984).
5. Brinster, R. L. *et al.* *Cell* **37**, 367–379 (1984).
6. Jenkins, N. A. & Copeland, N. G. in *Important Advances in Oncology* (eds DeVita, V. T., Hellum, S. & Rosenberg, S. A.) 61–77 (Lippincott, Philadelphia, 1989).
7. Hanahan, D. *Nature* **315**, 115–122 (1985).
8. Mahon, K. A. *et al.* *Science* **235**, 1622–1628 (1987).

9. Baetge, E. E., Behringer, R. R., Messing, A., Brinster, R. L. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3648–3652 (1988).
10. Botteri, F. M., van der Putten, H., Wong, D. F., Sauvage, C. A. & Evans, R. M. *Molec. cell. Biol.* **7**, 3178–3184 (1987).
11. Ts'o, M. O. M., Zimmerman, L. E. & Fine, B. S. *Am. J. Ophthalmol.* **69**, 339–349 (1970).
12. Ts'o, M. O. M., Fine, B. S. & Zimmerman, L. E. *Am. J. Ophthalmol.* **69**, 350–359 (1970).
13. Albert, D. M., Sang, D. N. & Craft, J. L. *Jap. J. Ophthalmol.* **22**, 358–374 (1978).
14. Flexner, S. *Bull. Johns Hopkins Hosp.* **2**, 115–119 (1891).
15. Wintersteiner, H. *Das Neuroepithelioma Retinae: Eine Anatomische und Klinische Studie* (Franz Deuticke, Vienna, 1887).
16. Wright, J. H. *J. exp. Med.* **12**, 556 (1910).
17. Kyritsis, A. P., Tsokos, M., Triche, T. J. & Chader, G. J. *Nature* **307**, 471–473 (1984).
18. Kivela, T. *Antigenic Properties of Retinoblastoma Tissue in Relation to Neuronal and Glial Cells of the Human Retina* 51–67 (University of Helsinki, 1987).
19. Stone, J. C. *et al.* *Genomics* **5**, 70–75 (1989).
20. Hsieh, C.-L. *et al.* *Somatic Cell molec. Genet.* **15**, 461–464 (1989).
21. Whyte, P. *et al.* *Nature* **334**, 124–129 (1988).
22. Dyson, N., Howley, P. M., Munger, K. & Harlow, E. *Science* **243**, 934–936 (1989).
23. Ludlow, J. W. *et al.* *Cell* **56**, 57–65 (1989).
24. Green, M. R. *Cell* **56**, 1–3 (1989).
25. Bernards, R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 6474–6478 (1989).
26. Harlow, E. *et al.* *J. Virol.* **39**, 861–869 (1981).
27. Lane, D. P. & Crawford, L. V. *Nature* **278**, 261–263 (1979).
28. Sarnow, P., Yo, Y., Williams, J. & Levine, A. J. *Cell* **28**, 387–394 (1982).
29. Efrat, S., Baekkeskov, S., Lane, D. & Hanahan, D. *EMBO J.* **6**, 2699–2704 (1987).
30. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
31. Mellon, P. L., Parker, V., Gluzman, Y. & Maniatis, T. *Cell* **27**, 279–288 (1981).
32. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
33. Toozé, J. (ed.) *Molecular Biology of Tumor Viruses, Part 2* (Cold Spring Harbor Laboratory, New York, 1981).
34. Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1986).
35. Tedder, T. F. *et al.* *J. Immunol.* **141**, 4388–4394 (1988).

ACKNOWLEDGEMENTS. We thank R. Weiner and T. Dryja for discussions, and R. Davis for preliminary histological analysis. This work was supported by the NIH (P.L.M., D.M.A., J.M.O., D.M. and C.M.D.), the Edward J. Mallinckrodt Foundation and the Searle Scholarship Foundation (R.B.), the American Cancer Society and the J. Aron Charitable Foundation (J.J.W.) and the Massachusetts Lions Eye Research Fund (J.M.O. and D.M.).

Non-hydrophobic extracytoplasmic determinant of stop transfer in the prion protein

C. Spencer Yost^{*†}, Charles D. Lopez[†], Stanley B. Prusiner^{‡§}, Richard M. Myers^{‡§} & Vishwanath R. Lingappa^{†||}

Departments of * Anesthesia, † Physiology, ‡ Neurology, § Biochemistry and Biophysics, and || Medicine, University of California, San Francisco, California 94143, USA

A UNIVERSAL feature of integral transmembrane proteins is a hydrophobic peptide segment that spans the lipid bilayer. These hydrophobic domains are important for terminating the translocation of the polypeptide chain across the membrane of the endoplasmic reticulum (a process termed stop transfer) and for integrating the protein into the bilayer^{1–5}. But a role for extracytoplasmic sequences in stop transfer and transmembrane integration has not previously been shown. Recently, a sequence which directs an unusual mode of stop transfer has been identified in the prion protein. This brain glycoprotein exists in two isoforms^{6–8}, which are identical both in primary amino-acid sequence and in containing phosphatidylinositol glycolipid linkages at their C termini, which can be cleaved by a phosphatidylinositol-specific phospholipase C⁹. But only one of the isoforms (PrP^C) is released from cells on treatment with this phospholipase¹⁰, indicating that the two isoforms have either different subcellular locations or transmembrane orientations. Consistent with this is the observation of two different topological forms in cell-free systems^{11,12}. An unusual topogenic sequence in the prion protein seems to direct these alternative topologies (manuscript in preparation). In the wheat-germ translation system, this sequence directs nascent chains to a transmembrane orientation; by contrast, in the rabbit reticulocyte lysate system, this sequence fails to cause stop transfer of most nascent chains. We have now investigated determinants in this unusual topogenic sequence that direct transmembrane topology, and have demonstrated that (1) a lumenally disposed

charged domain is required for stop transfer at the adjacent hydrophobic domain, (2) a precise spatial relationship between these domains is essential for efficient stop transfer, and (3) codons encompassing this hydrophilic extracytoplasmic domain confer transmembrane topology to a heterologous protein when engineered adjacent to the codons for a normally translocated hydrophobic domain. These results identify an unexpected functional domain for stop transfer in the prion protein and have implications for the mechanism of membrane protein biogenesis.

We constructed a series of deletions throughout the prion protein (PrP) to demonstrate the significance of the unusual topogenic sequence. Figure 1 illustrates three of these mutants and compares them with the wild-type (wt) PrP molecule. We expressed these PrP mutants in the wheat-germ (WG) translation system supplemented with dog pancreas rough microsomal membranes. We digested aliquots with proteinase K in the presence or absence of the non-ionic detergent Triton X-100, and subjected them to SDS-PAGE. The protein encoded by the mutant PrP ha 32 (Fig. 2a, lanes 1–4) had the same topology as wt PrP (PrP ha 1; Fig. 2b, lanes 5–8), as determined by the size, intensity and proportion of the characteristic N-terminal transmembrane fragment generated on proteolysis of intact membranes (compare Fig. 2b, lanes 7 and 8 with Fig. 2a, lanes 3 and 4)¹². Proteolysis in the presence of detergent completely hydrolysed translation products for PrP ha 32 and all other PrP mutants (Figs 2, 3 and 5). These results were as expected, in view of the large body of evidence indicating that only very discrete regions of a protein are involved in determining topology^{13,14}. Indeed, deletion of other regions in the C-terminal end of PrP did not affect the transmembrane topology (data not shown).

PrP mutant ha 30 (Fig. 2a, lanes 5–8) which has deleted transmembrane regions has, as predicted, a fully translocated topology, as demonstrated by glycosylation and protection from protease digestion in the absence of detergents. This is consistent with previous studies demonstrating an absolute requirement for the hydrophobic domain in transmembrane integration^{1,2,15}.

The gene encoding the mutant PrP ha 28 (Fig. 1) has a deletion of 24 codons just 5' to the region encoding the first transmembrane region, but leaves the entire transmembrane region

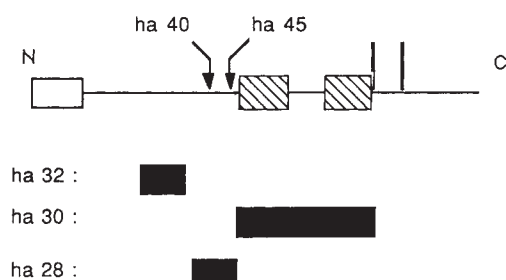


FIG. 1 Deletion mutagenesis of PrP ha 1. The 254-amino-acid residues of Syrian hamster PrP are shown at the top. Open box indicates 22-residue signal peptide. Matched boxes show membrane-spanning regions. Vertical bars indicate possible N-linked glycosylation sites. Arrows indicate positions at which insertions generating PrP ha 40 and PrP ha 45 were made. Black boxes below indicate the regions of PrP ha 1 that were deleted to generate each of the indicated mutants, PrP ha 32, PrP ha 30 and PrP ha 28, corresponding to the deletion of codons 39–62, 91–158 and 63–87, respectively. METHODS. Plasmids were constructed as follows. Plasmid pSP PrP ha 1: PrP complementary DNA from Syrian hamster isolate was cloned into pSP64T (refs 34, 35). Plasmid pSP PrP ha 28: a *Kpn*I site was introduced by site-directed mutagenesis at codon 73 of PrP ha 1 construct to generate plasmid pSP PrP ha 23. Plasmid was opened at *Kpn*I site, treated sequen-

intact. Surprisingly, this deletion had a dramatic effect on transmembrane topology. Many chains were glycosylated and fully protected from protease, with few chains displaying the characteristic transmembrane topology expected for wt PrP (Fig. 2b, compare lane 3 with lane 7). We noted that the fraction of chains glycosylated varied with each mutant. Experiments in which N-linked glycosylation was abolished demonstrated that glycosylation was not a factor in determining transmembrane or secretory topology of either PrP or its deletion mutants (data not shown; see legend to Fig. 2). The region deleted in PrP ha 28 has a strikingly positive charge distribution, containing four

tylally with nuclease *Bal*/31, Klenow fragment of DNA polymerase 1 and T4 DNA ligase. Plasmid pSP PrP ha 30: plasmid pSP PrP ha 1 was cut with *Nae*I in the presence of ethidium bromide, recut with *Hinc*2 and closed with T4 DNA ligase. Plasmid pSP PrP ha 32: plasmid pSP PrP ha 1 was cut with *Nco*I, a 270-base pair (bp) fragment isolated and religated back into the vector *Nco*I site. Plasmid pSP PrP ha 40: 350-bp *Eco*R1–*Bst*E2 fragment encoding chimpanzee globin residues 1–110, and including a previously inserted sequence encoding an N-linked glycosylation site, was generated from plasmid pSP SLSTgG¹⁷. After treatment with Klenow fragment of DNA polymerase 1, *Kpn*I 8mer linkers were added to the blunt ends and the fragment was religated into the *Kpn*I site of plasmid pSP PrP ha 23. Plasmid pSP PrP ha 45: *Bam*H1 10mer linkers were added to the same blunted fragment from plasmid pSP SLSTgG described above treated with *Bam* methylase. An upstream *Nae*I site in a noncoding region of plasmid pSP PrP ha 1 was eliminated by digestion with *Nhe*I and *Sph*I, and treated with Klenow fragment of DNA polymerase 1, mung bean nuclease, and T4 ligase. The sole remaining *Nae*I site located in the coding region of PrP ha 1 was cut with *Nae*I, and *Bam*H1 10mer linkers added. The plasmid was then cut with *Bam*H1, and the *Bam*H1-tailed fragment inserted. All constructions involving *Bal*/31 treatment were sequenced to determine digestion end points. All constructions were expressed by transcription-linked translation and the encoded product confirmed by immunoreactivity to relevant antisera (monospecific anti-peptide sera to sequences in the N- and C-terminal domains of PrP or to anti-human globin serum), and the expected size confirmed by SDS-PAGE.

lysine and two histidine residues and completely lacking hydrophobic character (Fig. 4a). The behaviour of this mutant is consistent with the hypothesis that this hydrophilic extracytoplasmic region (from here on termed STE for stop transfer effector) promotes the integration of the adjacent hydrophobic sequence into the bilayer.

To understand better the spatial relationship of STE to the hydrophobic domain, we inserted codons from the cytosolic protein α -globin into the coding region of PrP. Because α -globin contains no information for either start or stop transfer^{16–18}, it served as a 'spacer' and allowed us to determine the importance

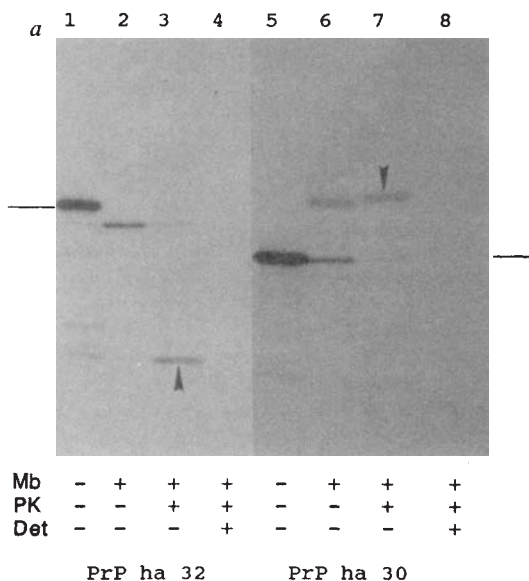
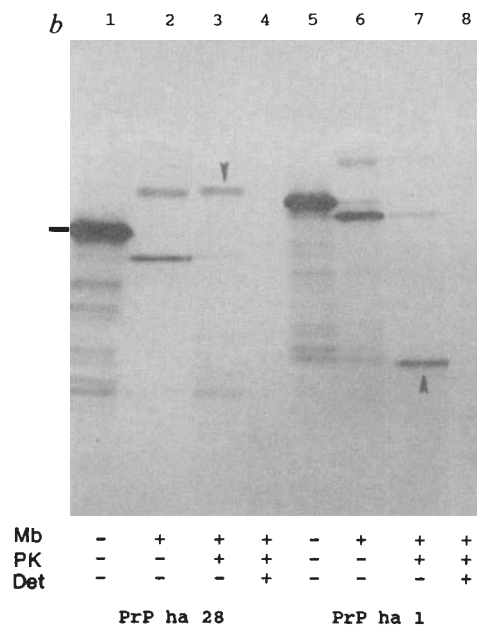


FIG. 2 Cell-free transcription-linked translation-coupled translocation of products encoded in plasmids pSP PrP ha 1 and deletion mutants described in Fig. 1 analysed by SDS-PAGE and fluorography of total products as previously described¹⁹. Transcriptions were performed using SP6 polymerase³⁴. Translations used WG extracts with dog pancreas microsomes (2.5 A_{280} units per ml) where indicated, as previously described¹². Aliquots post-translationally proteolysed with or without detergent are indicated below each panel. a, PrP ha 32 (lanes 1–4) and PrP ha 30 (lanes 5–8); b, PrP ha 28 (lanes 1–4) and PrP ha 1 (lanes 5–8). In both a and b, lanes 1 and 5 are products in the absence of membranes and lanes 2 and 6 are products in the presence of membranes; in lanes 3 and 7, the products synthesized in the presence of membranes were digested with proteinase K (PK); in



lanes 4 and 8 proteolysis was in the presence of detergent. Arrows denote the main protected products after proteolysis. All proteolysis was performed at 4 °C for 1 h with 0.2 mg ml⁻¹ PK, 10 mM CaCl₂, 10 mM Tris-HCl buffer, pH 8.0, and with or without 1% Triton X-100. PK digestions were terminated by adding phenylmethylsulphonyl fluoride to 1 mM in dimethyl sulphoxide and boiling in 1% SDS–0.1 M Tris-HCl, pH 8.9, for 5 min. Experiments using an acceptor tripeptide to inhibit N-linked glycosylation showed that glycosylation was not a determinant of transmembrane or secretory topology³⁶. These conclusions were confirmed by quantitative densitometry adjusted for methionines (data not shown). Abbreviations: Mb, membrane; PK, proteinase K; Det, detergent.

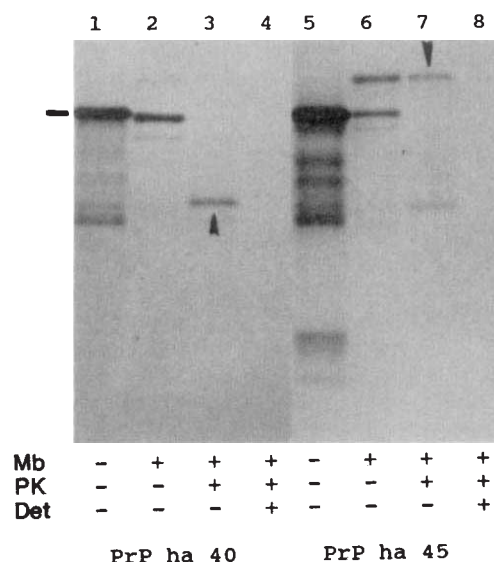


FIG. 3 Cell-free transcription-linked translation-coupled translocation of products encoded in plasmids pSP PrP ha 40 and 45 as described in Fig. 1 and analysed by SDS-PAGE and fluorography of total products. Transcription, translation and proteolysis are as described in Fig. 2. PrP ha 40 and PrP ha 45 translated without membranes (lanes 1 and 5), with membranes (lanes 2 and 6), and with membranes followed by PK digestion in the absence (lanes 3 and 7) or presence (lanes 4 and 8) of detergent. Arrows denote the main protected products. Abbreviations, as for Fig. 2.

of location for the function of STE. Codons 1-110 of the chimpanzee α -globin gene, as well as an eight-codon sequence encoding an *N*-linked glycosylation site¹⁹, were inserted either just *N*-terminal to the sequence encoding STE (PrP ha 40), or between the sequences encoding STE and the hydrophobic transmembrane region (PrP ha 45), as indicated by arrows in Fig. 1. The fusion protein PrP ha 40 had the same topology as wt PrP ha 1 (Fig. 3, lanes 1-4 compared with Fig. 2b, lanes 5-8). But PrP ha 45 had mainly the topology of deletion mutant PrP ha 28 (Fig. 3, lanes 5-8 compares with Fig. 2, lanes 1-4). Globin residues placed between STE and the hydrophobic transmembrane domain therefore disrupt stop transfer. Apparently the spatial relationship of STE to the hydrophobic domain is critical for transmembrane integration of PrP.

The hydrophathy profile of PrP ha 1 and the mutants PrP ha 28, ha 40 and ha 45 using the consensus program of Eisenberg²⁰ are shown in Fig. 4. The amino-acid sequences of the relevant regions of these molecules are also listed. The plot shows that the disruptive effect of the deletion PrP ha 28 and of the insertion in PrP ha 45 are not a result of alteration in hydrophobicity or introduction of specific sequences. Rather, the observed effects seem to be a consequence of abolishing the functional relationship between STE and the hydrophobic anchor.

Finally, we determined whether STE could induce stop transfer at a normally translocated hydrophobic domain, and if so, whether these observations could be extended to stop transfer in living cells. To do this we took advantage of earlier observations with chimaeric proteins¹⁷. In *Xenopus* oocytes, translocation initiated by an *N*-terminal signal sequence cannot be terminated by another signal sequence that follows internally

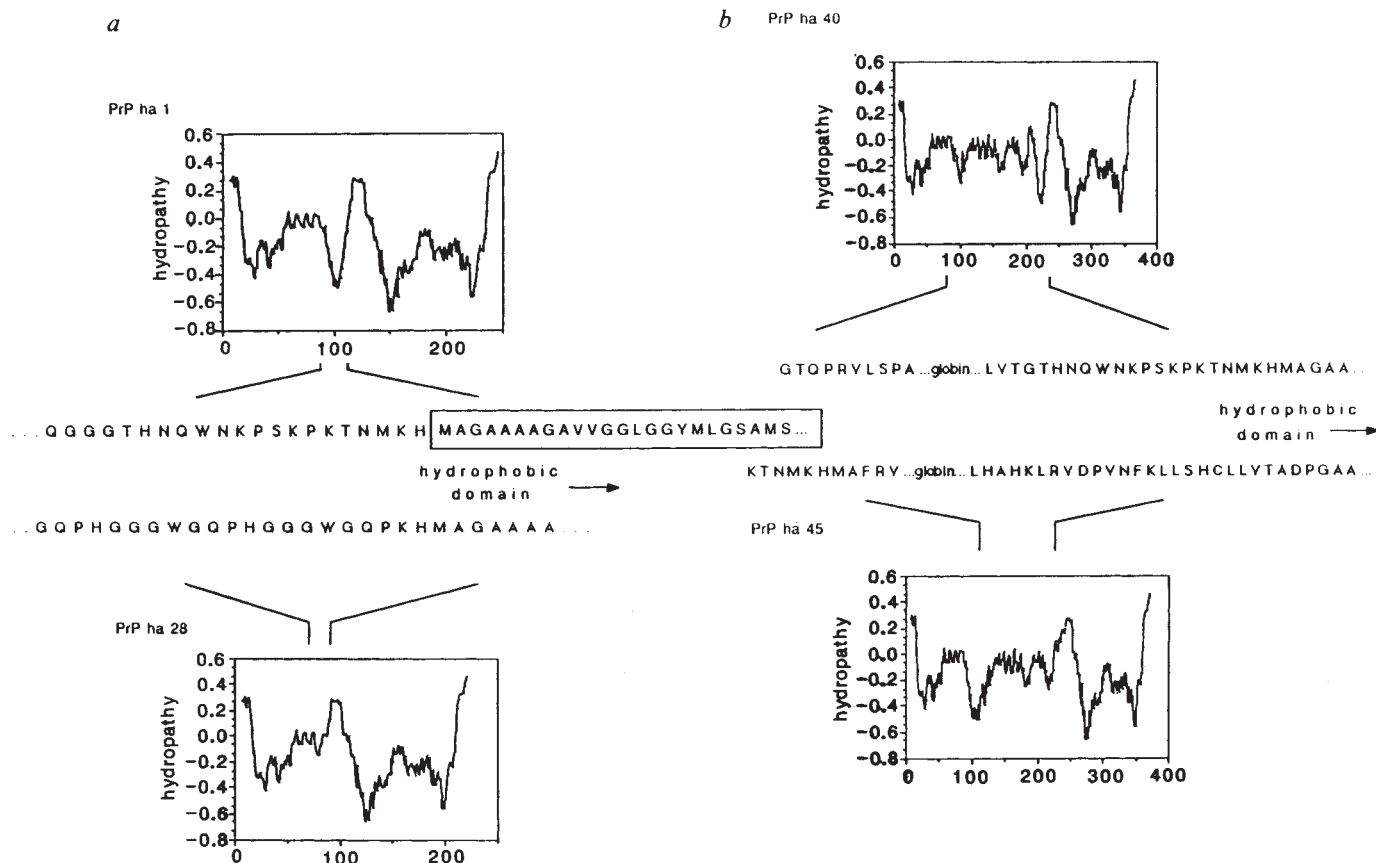


FIG. 4 Hydrophathy plots for PrP ha 1, PrP ha 28, PrP ha 40 and PrP ha 45. *a*, Upper plot shows Eisenberg consensus profile (14-residue window) of Syrian hamster PrP. The single amino-acid code for the hydrophilic region preceding the transmembrane segment is listed below. Lower plot shows the profile for the 230-amino-acid sequence of pPrP ha 28 in which the hydrophilic domain has been deleted. Listed above is the amino-acid sequence across the deletion junction as determined by dideoxy-chain-termination sequencing

of plasmid DNA. Boxed residues show the full sequence of the transmembrane domain. *b*, Hydrophathy profile of PrP ha 40 and PrP ha 45. Upper panel shows plot of PrP ha 1 with globin inserted in-frame before STE, with the amino-acid sequence of points of globin-PrP fusion shown below. Lower panel shows plot of PrP ha 1 with globin inserted between STE and the hydrophobic domain. Sequence above shows points of globin-PrP fusion. Abbreviations, as for Fig. 2.

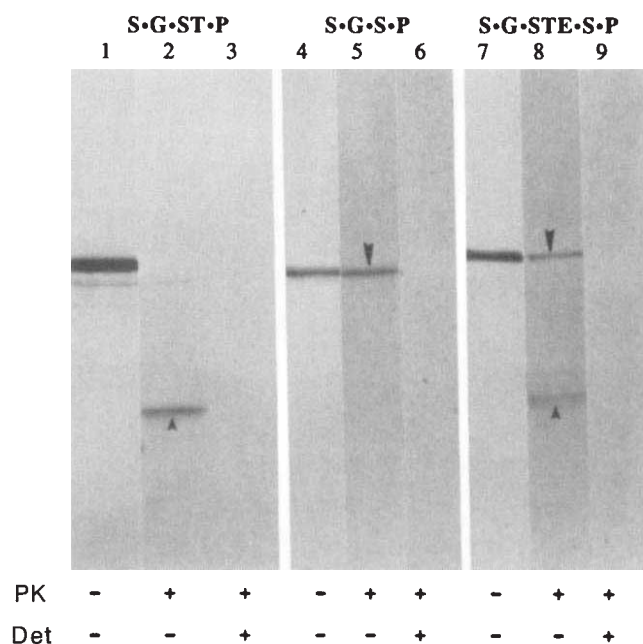


FIG. 5 Expression of chimaeric proteins in *Xenopus* oocytes. Chimaeric proteins were engineered (see ref. 17, and below) and expressed in stage VI *Xenopus* oocytes by microinjection of RNA transcripts and [35 S]-methionine as previously described¹⁷. Oocytes were labelled for 4 h, homogenized in 0.25 M sucrose, 0.05 M Tris buffer pH 7.6, 100 mM potassium acetate, 5 mM magnesium acetate, 5 mM magnesium acetate, 1 mM DTT, and an aliquot of homogenate proteolysed with proteinase K in the presence and absence of detergent as previously described¹⁷. Products were then immunoprecipitated with antisera raised against human globin or with non-immune serum (data not shown) and analysed by SDS-PAGE. Arrows pointing downwards denote fully protected secreted products; arrows pointing upwards denote protease-protected fragments from transmembrane chains. Plasmid pSP S.G.STE.S.P was engineered as follows. Plasmid pSP S.G.S.P was cleaved with restriction endonuclease *Bst*II, treated with Klenow fragment of DNA polymerase I, and ligated to *Xba*I 12mer linkers. The resulting intermediate plasmid was cleaved with *Xba*I, treated with calf intestinal alkaline phosphatase and ligated in the presence of an excess of a fragment encompassing the 90 bp of the PrP encoding region from *Nae*I to *Nco*I sites generated from intermediates by treatment with Klenow fragment, addition of *Xba*I 8mer linkers, cleavage with *Xba*I and purification by agarose gel electrophoresis.

(construct S.G.S.P, so termed to indicate component protein domains; see legend to Fig. 5). By contrast, a genuine stop transfer sequence in the same context results in transmembrane topology (S.G.ST.P). Thus, we engineered STE immediately next to the internal signal sequence of the former chimaera and expressed the new hybrid (termed S.G.STE.S.P) in *Xenopus* oocytes. If STE can confer stop-transfer activity it would be expected to terminate chain translocation at the normally translocated hydrophobic internal signal sequence. This would generate a protected globin reactive fragment on proteolysis of intact membrane vesicles, similar to that observed for S.G.ST.P. Figure 5 shows the topology of S.G.STE.S.P and of the original chimaeric proteins that we used as controls. The STE domain conferred transmembrane topology to 50% of the chains (as determined by quantitative densitometry corrected for methionine distribution), in contrast to the behaviour of S.G.S.P chains, which were completely translocated (see lane 7). This percentage is similar to that observed for other chimaeras in which STE is next to the hydrophobic first transmembrane region (TM1) of PrP (C.D.L., C.S.Y., S.B.P., R.M.M. and V.R.L., manuscript in preparation).

Because of the requirement for thermodynamic stability of proteins in membranes, it is not surprising that hydrophobic domains are essential for transmembrane integration^{1-5,21,22}. But a role for extracytoplasmic sequences in stop transfer, although previously suggested^{13,16}, has not been demonstrated. The

studies presented here indicate such a role for STE in PrP biogenesis. In view of the increasing evidence that the initiation of translocation across membranes is mediated by receptors²³, and that translocation itself occurs through an aqueous channel²⁴, it might be expected that stop transfer has some receptor-mediated features. By analogy with other protein-mediated systems, in some cases the efficiency of stop transfer might depend on regulatory factors in the cytosol or the membrane, or in both. The seemingly 'inefficient' stop transfer activity of PrP is independent of ongoing protein synthesis and dependent on cytosolic factors (C.D.L., C.S.Y., S.B.P., R.M.M. and V.R.L., manuscript in preparation). We suggest here that STE operates as a ligand for such receptor-mediated stop transfer.

Studies of other integral transmembrane proteins are necessary to determine whether these findings are a universal feature of transmembrane protein biogenesis or an unusual feature limited to a small class of proteins including PrP. The corresponding extracytoplasmic regions of other transmembrane proteins are quite divergent in sequence and charge. Also, the hydrophobic domains of various native and engineered transmembrane proteins differ significantly in their overall hydrophobicity as well as in the specific amino-acid residues they contain²⁵⁻²⁸. It is noteworthy that two positively charged residues are retained immediately before the transmembrane region in the deletion mutant PrP ha 28. The observed effect of STE is, therefore, probably not specifically mediated by charge but rather by more subtle structural features, as is likely to be the case for other receptor-ligand interactions involving chain translocation^{29,30}.

The PrP^{Sc} isoform of PrP is a component of the infectious particle that causes scrapie, a degenerative neurological disease of animals which is related to similar diseases in humans⁸. The PrP protein from mouse strains with different scrapie incubation times, as well as an isoform of human PrP linked to inherited prion disease, Gerstmann-Strausler syndrome, all have unique substitutions of non-charged amino acids in STE³¹⁻³³. Perhaps control over the action of STE is important for the generation of both secreted and membrane-bound forms of PrP (refs 11, 12), for the as-yet-unknown function of PrP in normal brain, or for the generation of PrP^{Sc} during scrapie. □

Received 16 November; accepted 13 December 1989.

- Rose, J. K. & Bergmann, J. E. *Cell* **30**, 753-762 (1982).
- Gething, M. J. & Sambrook, J. *Nature* **300**, 598-603 (1982).
- Cutler, D. F. et al. *J. Cell Biol.* **102**, 902-910 (1986).
- Doyle, C. et al. *J. Cell Biol.* **103**, 1193-1204 (1986).
- Adams, G. A. & Rose, J. K. *Cell* **41**, 1007-1015 (1985).
- Oesch, B. et al. *Cell* **40**, 735-746 (1985).
- Meyer, R. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2310-2314 (1986).
- Prusiner, S. B. *New Engl. J. Med.* **317**, 1571-1581 (1987).
- Stahl, N. et al. *Cell* **51**, 229-240 (1987).
- Stahl, N. et al. *FASEB J.* **2**, A989, (1988).
- Hay, B. et al. *Biochemistry* **26**, 8110-8115 (1987).
- Hay, B. et al. *Molec. cell. Biol.* **7**, 914-920 (1987).
- Blöbel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496-1500 (1980).
- Perara, E. & Lingappa, V. R. in *Protein Transfer and Organelle Biogenesis* (eds Das, R.C. & Robbins, P.) 3-47 (Academic, San Diego, 1989).
- Boeke, J. D. & Model, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5200-5204 (1982).
- Yost, C. S. et al. *Cell* **34**, 759-766 (1983).
- Rothman, R. E. et al. *J. Biol. Chem.* **263**, 10470-10480 (1988).
- Lingappa, V. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 456-460 (1984).
- Perara, E. & Lingappa, V. R. *J. Cell Biol.* **101**, 2292-2301 (1985).
- Eisenberg, D. A. *Rev. Biochem.* **53**, 595-623 (1984).
- Davis, N. G. & Model, P. *Cell* **41**, 607-614 (1985).
- Zerial, M. et al. *Cell* **48**, 147-155 (1987).
- Widemann, M. et al. *Nature* **328**, 830-833 (1987).
- Gilmore, R. & Blöbel, G. *Cell* **42**, 497-505 (1985).
- Sivasubramanian, N. & Nayak, D. P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1-5 (1987).
- Gracia, G. L. & Hunter, E. *J. Cell Biol.* **105**, 1191-1203 (1987).
- Szczesna, S. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 738-742 (1988).
- Lipp, J. & Dobberstein, B. *J. Cell Biol.* **106**, 1813-1820 (1988).
- Siegel, V. & Walter, P. *Cell* **52**, 39-49 (1988).
- Lingappa, V. R. *J. clin. Invest.* **83**, 739-751 (1989).
- Carlson, G. E. et al. *Cell* **46**, 503-511 (1986).
- Hsiao, K. et al. *Nature* **338**, 342-345 (1989).
- Westaway, D. et al. *Cell* **51**, 651-662 (1987).
- Melton, D. et al. *Nucleic Acids Res.* **12**, 7035-7056 (1984).
- Basler, K. et al. *Cell* **46**, 417-428 (1986).
- Hansen, W. et al. *Cell* **45**, 397-406 (1986).

ACKNOWLEDGEMENTS. We thank Pablo Garcia for help with the hydropathy analysis program and members of our laboratories for critical reading of the manuscript. This work was supported by the NIH.

Software for smart users

This week's focus on scientific software features a program for protein and DNA sequence analysis, a 3-D graphics program and software to ease your way through complex mathematical problems.

THE Laboratory of Fluorescence Dynamics at the University of Illinois has a new software package—Globals Unlimited—designed for **time-resolved fluorescence decay analysis** in either the time or frequency domain (*Reader Service No. 100*). Multidimensional data surfaces that correspond to multiple experimental axes—temperature, wavelength, pressure and pH—can be analysed simultaneously, says the university. Analysis options include multi-exponential decays, multimodal distributed decay, excited state reaction, energy transfer quenching, time-resolved spectra and anisotropy decay. Globals Unlimited runs on any IBM PC or compatible and is priced at \$1,300 (US).

Image Analyst is a Macintosh-based **image processing software package** from Automatrix (*Reader Service No. 101*). It is designed to process data from video frame grabbers or from disc-stored images in PICT or TIFF format. Using an interactive tool palette, users can define a region of interest as small as a single pixel or covering the entire image. Over 200 types of measurement can be extracted from the image, says Automatrix. Image

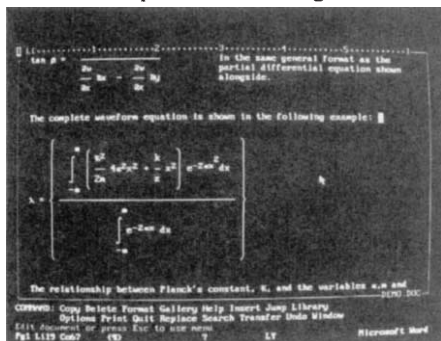


Image Analyst: Mac-based measurement software from Automatrix.

Analyst can be used for image enhancement, counting objects, determining density, shape, size, position and movement, and for conducting textural analyses. Three programs are available: Image Analyst 7.1, Image Analyst/Source and 'C' libraries. Image Analyst 7.1 is for users who do not want to write their own programs, and is configured for quick image processing and measurement routines. For repetitious applications, users can build sequences containing multiple measurements, which are executed by a single command. With Image Analyst/

Source the user can modify Image Analyst software, change an image's presentation, access additional algorithms and measurement functions, add custom algorithms, and interface to other I.O. devices. 'C' libraries are for users who need to solve complex and time-critical applications.

MathWord and ChemWord are two **scientific word processing programs** from Laboratory Software Ltd, which are fully integrated with Microsoft Word (*Reader Service No. 102*). MathWord provides a full Greek alphabet and a range of mathe-



Use MathWord for mathematical word processing—new from Laboratory Software.

matical and engineering symbols. ChemWord has all the facilities to produce chemical formulae. Mathematical equations and chemical formulae appear on the screen as they will be printed, and can be copied, re-positioned and edited while in the Word document. Parts of formulae can be 'lifted' out of diagrams and placed in either the body text or stored in the Word glossary. MathWord and ChemWord run on an IBM PC or compatible and support dot matrix and laserprinters, including PostScript.

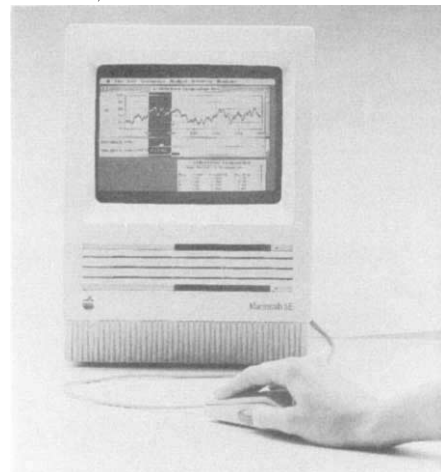
MultiCalc is Pharmacia Wallac Oy's latest **software for all types of label assays** (*Reader Service No. 103*). The software has data logging, processing, results calculation, data presentation and quality control capabilities, and users can customize the computer system to suit laboratory needs. MultiCalc is suitable for RIAs and IRMAs, as well as non-radioactive techniques such as enzyme, fluorescent, and luminescent label assays. In addition, MultiCalc can handle assays in both sequential and plate format, and supports PC networking facilities. Other applications include microbiological screening, receptor assays, multiple-isotope and sample dilution studies, kinetic and two-point measurements. MultiCalc runs on

any PC, XT, AT, PS/2 or compatible computer.

The Cryobase software program from Cymbus Bioscience is designed to **maximize storage space in cryogenic freezers** and streamline management methods at the same time (*Reader Service No. 104*). Cryobase has been written specifically to allow the user to record and cross reference information on stored items, while making the best use of the space available. This eliminates, says Cymbus, laborious handwritten documentation, possible inaccuracies in record keeping which can arise with a multi-user facility, and excessive liquid nitrogen consumption when searching for samples that are misplaced. Each location within the tray is identified by container number, file number, tray number, and position within the tray. Users also enter a short description of the sample. Once stored, the information can, for example, be used to locate empty spaces, search for a specific vial, and display the total contents of a tray. The \$300 (US) Cryobase program runs on any IBM PC or compatible.

Sequencing software

Macintosh users may be interested in the new **protein and DNA sequence analysis software** called MacVector from International Biotechnologies (*Reader Service No. 105*). MacVector allows the user to



MacVector—for in-depth DNA and protein sequence analysis.

generate circular plasmid maps, enter sequence data via a sonic digitiser, display multiple windows for in-depth data analysis, and perform sequence editing using the Mac edit menu. In addition, users can predict secondary protein structure,

perform hydrophilicity calculations and make antigenic determinations. Data can be shared with sequence data from other analysis systems such as GENEBANK, BIONET, UWGCG, text and Staden, without the need for reformatting, says IBI. Publication-quality graphics and slides, compatible with Macintosh word processing and desktop publishing programs, can be generated as TEXT or PICT files. MacVector sells for \$2,495 (US).

IntelliGenetics is offering new ProtPlot software for producing **plots of protein secondary structure** (*Reader Service No. 106*). Using ProtPlot, secondary structure, hydrophobicity and hydrophilicity, amphipathic α -helix and β -sheet, and hydrophobic moment analysis can be determined. ProtPlot accepts protein sequence files or secondary structure files as input. It can be used to analyse protein sequences up to 3,000 amino acids long, and can read single-letter protein sequence files entered by using a standard text editor or written by PC/GENE — IntelliGenetics' IBM PC-based system for sequence analysis. Output can be directed to printers, plotters or tabular files. Plots are output as schematic representations of overall protein secondary structure, or as helical wheel diagrams of polypeptide regions showing the spatial juxtaposition of side chains. In addition, putative glycosylation sites can be located and identified. ProtPlot runs on an IBM PC or compatible and costs \$400 (US and Canada) — \$250 (US and Canada) for PC/GENE owners.

Quantitation software

Biological Vision announces that the new **PCR quantitation applications software** for autoradiographic, photographic or colourimetric reporter systems is now available as part of the basic BVI 4000 DNA Vision System (*Reader Service No. 107*). The software can be used to quantitate the PCR product in any format — films, gels and microtitre plates. Its unique algorithms can extend the dynamic range of film by over two logs, says BVI.



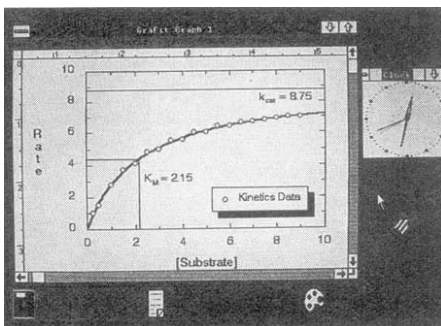
Biological Vision's PCR Quantitation software is now part of the basic BVI 4000 DNA image analysis system.

The \$27,500 (US) BVI DNA Vision System includes PCR quantitation and blot analysis applications programs.

Pædia Corporation has introduced **1-D densitometry software** for electrophoresis which can analyse densitometric data from stained gels, blots and films (*Reader Service No. 108*). Data can be taken from any densitometer with analog output by a stand-alone, microprocessor-based analog-to-digital converter connected to the computer's serial port. Features of the package include data smoothing, baseline subtraction, scaling, automatic and manual peak identification, three methods of plot integration, plot overlay, molecular weight determination, and straight and curved-line fitting. The macro instruction set allows the computer to be programmed for unattended operation. Software packages for both the Macintosh and IBM (AT and PS/2) computers or compatibles are available. Included in the \$1,825 (US) price is the analog-to-digital converter, cables and instructions.

Graphing aids

GraFit is a scientific **curve-fitting program** from Erithacus Software which makes use of the Microsoft Windows graphic environment (*Reader Service No. 109*). The program, written by Dr R.J. Leatherbarrow, Imperial College, London, uses non-linear regression to fit data to one or

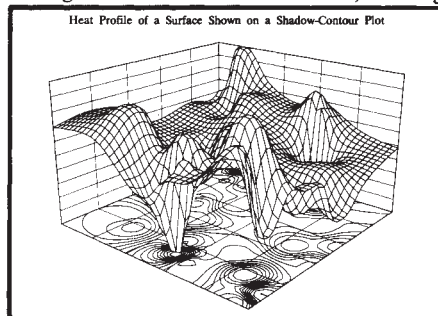


GraFit software fits curves to data.

more different equations. Fifty predefined equations are provided which cover a range of experimental situations, including pH activity profiles, dose-response curves, enzyme kinetics, ligand binding, enzyme inhibition, exponential analysis and radioimmunoassays. Further equations with up to ten unknown parameters can be added using the built-in equation editor. Graphs are created in a completely WYSIWYG manner, which allows for resizing and editing. Both axes can be logarithmically scaled, error bars can be added, graphs can include any number of overlaid data sets and/or calculated curves, and data can be exported into Windows Write and PageMaker. GraFit costs £290 (UK) and runs on a mouse-driven PC, AT or PS/2 compatible computer.

Graftool 2.1 is the latest version of 3-D Visions' interactive scientific and engineering **3-D graphics program** (*Reader Service No. 110*). Using the menu-driven

program, users can generate and analyse X-Y plots, bar, column and pie charts, polar graphs, Smith charts, parametric graphs, transfer function plots, surface plots, 3-D trajectory plots, data-eye diagrams, and one- and two-variable histograms in 3D. Graftool 2.1, costing



Manipulate data in 3D with Graftool.

\$495 (US), runs on IBM PCs, XT, AC or 386-compatibles and recognises data in ASCII, BASIC and DIF formats. On-screen data processing capabilities include contouring, cross-sectioning of 3-D surfaces, fast Fourier transforms, linear and non-linear regression and curve fitting. Graftool 2.1 features a unique 'data cursor' which allows the user to identify critical data points for on-screen analysis or data processing. All graphs and accompanying text can be moved, scaled and rotated in 3D. The zoom-and-pan function can enlarge any portion of the graph by fifteen orders of magnitude, says the company. Additional features include the facility for multi-coloured curves and surfaces, Greek and maths symbols, and 3D-to-2D data projections.

MultiFit Version 1.5 is a Macintosh-based **curve-fitting software package** from Day Computing (*Reader Service No. 111*). It offers the user a choice of non-linear fitting algorithms — Marquardt and Variable Metric — as well as complete control over non-linear fit calculation and convergence details, says Day. MultiFit 1.5 is supplied with a range of model equations, but new ones can be added containing up to eight variable and four constant parameters. As well as curve fitting, users can perform extrapolations and simulations. MultiFit 1.5 can also be used for simple and multiple linear regression, median fit, simple transformations, Deming model linear regression, descriptive statistics, and random number/sequence generation. Costing £80 (UK) or \$150 (US), MultiFit 1.5 is fully compatible with MultiFinder and comes with a 70-page manual explaining practical and theoretical details.

Fig. P is the latest **graphics software package** from Biosoft that runs on an IBM PC or PS/2 (*Reader Service No. 112*). Using Fig. P, users can employ various curve-fitting methods, including line plots, staircase plots, smooth linear or

polynomial least squares regression curve fit, free-form curvilinear plots, 1- and 2-D scattergrams, bargraphs and text-only charts. Fig. P accepts data from many laboratory instruments and up to 10,000 data points can be entered per figure, says Biosoft. Additional features include

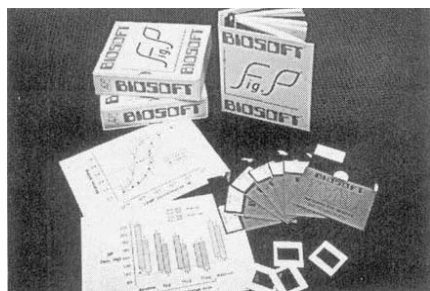
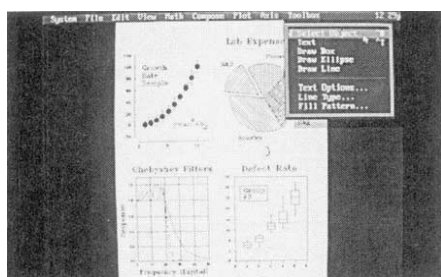


Fig. P: Biosoft's graphing program.

WYSIWYG screen presentation, Greek characters and scientific symbols, numerous line styles, multiple figures per page, and the automatic calculation of means, error/range bars. Graphs created using Fig. P, which costs \$399 (US) or £250 (UK), can be exported to word processing documents that use WordPerfect and desktop publishing programs such as PageMaker and Ventura. Output can be directed to laserprinters, slidemakers and plotters.

SigmaPlot 4.0 is the new version of Jandel Scientific's **graphing program** for IBM or compatible PCs (Reader Service No. 113). It offers a range of specialized types of graph, including scatter and line graphs, normal and stacked bar charts,



Jandel Scientific's SigmaPlot 4.0 plots a range of specialized graph forms.

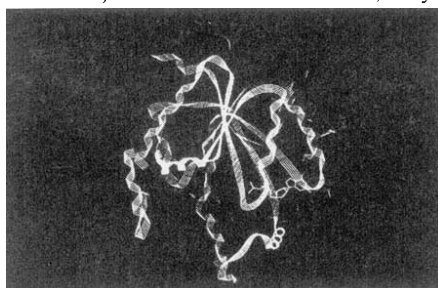
box plots and pie charts. Graphs can be presented with linear, log, semilog, log-log, natural log, probit, logit and probability scales. Error bars can be represented as standard error, standard deviation, 95 and 99 per cent confidence limits, with calculations based on either the arithmetic or geometric mean. Other features include a Greek alphabet, maths symbols, European language characters, and variable line thickness. SigmaPlot also provides full non-linear curve fitting whereby the user can define almost any equation — or sets of equations with up to 25 parameters and ten independent variables — and then fit the equation to their data. SigmaPlot costs \$495 (US) and is supplied with a 400-page manual.

Molecular modelling

ChemPolymer is the new Chem-X module for **polymer modelling** introduced by Chemical Design (Reader Service No. 114). The ChemPolymer toolkit exploits the open architecture and interactive analysis and display facilities of Chem-X, allowing users to study the potential of new polymers before they are produced in the laboratory. The software can be used to evaluate free volume, stiffness, elasticity and geometric properties of the polymer. Using techniques based on statistical thermodynamics, users can calculate the temperature dependence of properties such as entropies and free energies. ChemPolymer can be used to generate regular, blocked, and random copolymers from monomers, and can also handle branching to form network and star polymers. The software runs on DEC VAX computers and is fully integrated with the Chem-X molecular modelling system.

For researchers looking for ways to speed up the process of rational drug design, Molecular Design is making available Pergamon Press' *Comprehensive Medicinal Chemistry Drug Compendium* — CMC-3D — as an electronic database (Reader Service No. 115). CMC-3D is designed to be used in conjunction with MACCS-3D to examine the correlation between biological activity and geometric properties. It will include all six volumes of the *Drug Compendium* and will cover the rational design, mechanistic study and therapeutic applications of 5,400 compounds that have been marketed as drugs. Each entry will include information on chemical structure, generic name, CAS registry number, drug class, biological activity, literature references, originating company, physical properties and 3-D structure. CMC-3D will be available in March.

Tripos Associates has introduced SYBYL 5.3 — an integrated and extendable software kit for **3-D molecular modelling and analysis** (Reader Service No. 116). SYBYL 5.3 features, says



Show backbone configurations with SYBYL, 5.3.

Tripos, major enhancements to the biopolymer, polymer and dynamics add-on modules. In addition, the software offers improved graphical analysis and display tools, as well as various levels of menu

complexity. Faster and more efficient biopolymer research is supported by several new tools in SYBYL/Biopolymer: these include a protein loop search for constructing backbone configurations and protein homology modelling, and tools for manipulating macromolecules, including insertions, deletions, point mutations

ADVERTISEMENTS

CUSTOM-MADE & PURIFIED DNA PEPTIDES

\$8/base
50-100 µg
40 hour delivery

Quotes available
50-300 mg
1-3 wks delivery

GUARANTEED QUALITY & QUANTITY

V.G. CORP.

1200 Wonderland Rd. S.
Bldg #7, Unit #2
London, Ontario
Canada N6L 1A8

CALL OR FAX: (519) 652-3758

Anti-Vibration Tables, Patch
Clamp Amplifiers, Capillary
Glass, Epi-Fluorescence Systems,
Electrode Pullers Etc.,
FOR A FULL RANGE OF
ELECTROPHYSIOLOGICAL
PRODUCTS:

INTRACEL LTD

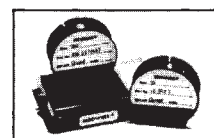
Unit 4, Station Road, Shepreth,
Royston Herts. UK SG8 6PZ
TEL: (0763) 262680

SPECIAL PRICE CUSTOM DNA

100 µg, 58.00 FF A BASE
FOR THE FIRST 15 BASES.
49.00 FF, each additional base
BIOPROBE SYSTEMS

CALL : 33-1-48-51-66-22
FAX : 33-1-48-51-59-90

Teflon* Coated and Bare Wire from MEDWIRE

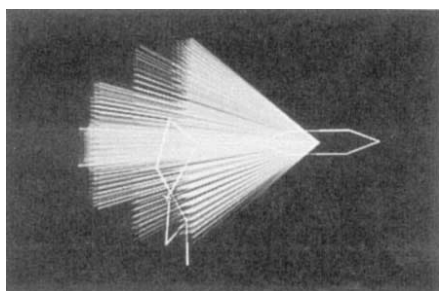


Medwire provides
the highest quality
platinum, platinum alloy, silver and stainless
steel wire in both single and multistrand
configurations.

For prompt delivery or additional
information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E.I. DuPont & Co.



3-D modelling with SYBYL 5.3.

and conformational analyses. SYBYL/Polymer's 'polymer build' facility provides an interactive means of constructing and modelling synthetic polymers from pre-built monomers or repeat units. It also allows the estimation of a variety of

ADVERTISEMENTS

ACTIVITY MONITOR

Automatic measurement of movement from the video (TV) image. For water maze (pool), open field, radial maze etc. IBM-PC compatible. Software for latency, path length, percent time moving, position preference etc. Options for rearing and social interaction.

In the U.S. call 1-800-225 9261

HVS IMAGE

18 Ormond Crescent

Hampton TW12 2TH

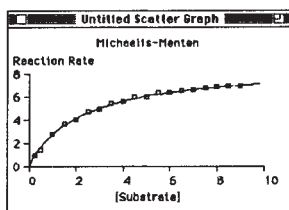
England

Telex: 8813487 LCLAD

Tel: +44 1 783 1152



Curve Fitting



MultiFit 1.5

for Apple Macintosh

Non-Linear Regression to User-Defined model

Simple, Multiple, Deming and Robust Linear Regression
Polynomial Regression
and much more...

£80 or US\$150

Orders and enquiries to:

DayComputing

P.O.Box 327, Milton,
Cambridge, CB4 4WL, UK

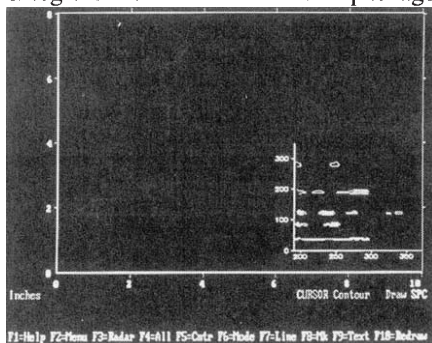
Fax +44 223 412709

physicochemical properties. SYBYL 5.3 will initially be available for DEC VAX platforms, but further releases are expected.

Dedicated programs

Philips has introduced a new **kinetics program** for users of the PU8750 Series of UV/visible spectrophotometers (*Reader Service No. 117*). The PU8713 program is based on Michaelis-Menton kinetics and will calculate K_m and V_{max} values from an interactive database of up to five substrate concentrations in triplicate and at constant enzyme concentrations. In addition, the results of two data sets can be compared to determine the effect of an inhibitor: K_i values are automatically calculated and displayed for competitive, non-competitive and uncompetitive inhibition. Other features of the software include automatic linearity checking, or 'linearity off' selection whereby multiple cycle results can be presented as $\Delta A \text{ min}^{-1}$ and rate units with their average values. There are three calculation methods to choose from: $1/V$ vs $1/S$, S/V vs S and V vs S . Using a PU8732 cell programmer to speed throughput, the PU8713 software can be configured for parallel kinetics for the simultaneous measurement of up to eight cells.

Analysing and presenting chromatographic data generated from Applied Biosystems' 1000S diode array detector need not be a chore with the company's new interactive software package called LabCalc (*Reader Service No. 118*). Designed to run on IBM PCs or selected compatible computers, LabCalc can be used to acquire, store, retrieve, process and display chromatographic data. LabCalc can perform 'standard' chromatographic calculations such as peak detection and integration, as well as interactive and automatic peak picking. Using Array Basic, users can create custom programs and integrate them into the LabCalc package.



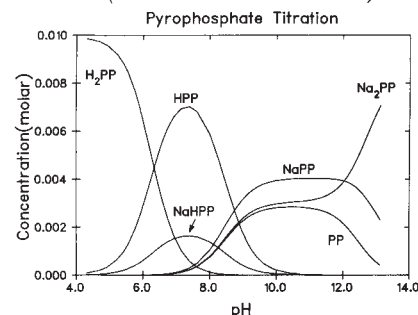
LabCalc chromatography software.

When in draw mode, users can overlay chromatograms and spectra, draw in axes, add labels and change fonts, producing, says Applied Biosystems, diagrams of presentation quality. Data can be exported to desktop publishing programs such as PageMaker and Ventura.

Problem solving

The collaborative efforts of Fraser Williams Scientific Systems and Preston Scientific have produced SpectraSearch — an **NMR spectral database** with search and display facilities (*Reader Service No. 119*). The user inputs a query chemical structure and the system carries out a structure and sub-structure search against a database of 8,000 organofluorine NMR spectra. Search results can be viewed along with full bibliographic information and spectral data, including both chemical shifts and coupling constants. The chemical shift search facility provides a means of identifying probable structures based on the actual chemical shift range obtained from the experimental NMR result. In addition to the fluorine database, Fraser Williams can provide a user-input module for creating a custom database. SpectraSearch runs on any PC or compatible and costs \$3,600 (US) for the first year — \$2,300 (US) for each subsequent year.

Save time and effort working out the solutions to difficult **aqueous chemical equilibrium problems** with EQUIL — the new program from MicroMath Scientific Software (*Reader Service No. 120*). The



EQUIL provides software solutions to mathematical problems.

user specifies the problem by defining solution properties such as pH and ionic strength. The program then automatically selects the appropriate equilibrium reactions from a built-in chemical equilibrium database and constructs mass-balance relationships with an 'equilibrium compiler'. Users can display output as pH titration curves, species concentration diagrams and complex formation curves. EQUIL can also determine activity coefficients calculated by a modification of the extended Debye-Huckel approach, and can indicate the degree of saturation with respect to each of the possible solid phases. EQUIL software has interactive graphics capabilities and output can be directed to plotters and printers. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.